## DOCTOR OF PHILOSOPHY

# Skin barrier dysfunction in common genetic disorders

Chen, Huijia

*Award date:*
2011

DOCTOR OF PHILOSOPHY

# Skin barrier dysfunction in common genetic disorders

Huijia Chen

2011

University of Dundee

# APPENDIX I

# PUBLICATIONS

***Vicky Gagnon[1,2], Danielle Larouche[1,2], Rémi Parenteau-Bareil[1], Marie Gingras[1], Lucie Germain[1] and François Berthod[1]***

*[1]Centre LOEX de l'Université Laval, Centre de Recherche FRSQ du Centre Hospitalier Affilié Universitaire de Québec and Département de Chirurgie, Faculté de Médecine, Université Laval, Québec City, Québec, Canada*
*E-mail: francois.berthod@chg.ulaval.ca*
*[2]These authors contributed equally to this work.*

**REFERENCES**

Blais M, Grenier M, Berthod F (2009) Improvement of nerve regeneration in tissue-engineered skin enriched with schwann cells. *J Invest Dermatol* 129:2895–900

Botchkarev VA, Eichmuller S, Johansson O *et al.* (1997) Hair cycle-dependent plasticity of skin and hair follicle innervation in normal murine skin. *J Comp Neurol* 386: 379–95

Caissie R, Gingras M, Champigny MF *et al.* (2006) *In vivo* enhancement of sensory perception recovery in a tissue-engineered skin enriched with laminin. *Biomaterials* 27:2988–93

Gingras M, Bergeron J, Dery J *et al.* (2003a) *In vitro* development of a tissue-engineered model of peripheral nerve regeneration to study neurite growth. *FASEB J* 17:2124–6

Gingras M, Paradis I, Berthod F (2003b) Nerve regeneration in a collagen-chitosan tissue-engineered skin transplanted on nude mice. *Biomaterials* 24:1653–61

Hamalainen HA, Warren S, Gardner EP (1985) Differential sensitivity to airpuffs on human hairy and glabrous skin. *Somatosens Res* 2:281–302

Hendrix S, Picker B, Liezmann C *et al.* (2008) Skin and hair follicle innervation in experimental models: a guide for the exact and reproducible evaluation of neuronal plasticity. *Exp Dermatol* 17:214–27

Larouche D, Cuffley K, Paquet C *et al.* (2011) Tissue-engineered skin preserving the potential of epithelial cells to differentiate into hair after grafting. *Tissue Eng Part A* 17:819–30

Peters EM, Botchkarev VA, Muller-Rover S *et al.* (2002) Developmental timing of hair follicle and dorsal skin innervation in mice. *J Comp Neurol* 448:28–52

Provitera V, Nolano M, Pagano A *et al.* (2007) Myelinated nerve endings in human skin. *Muscle Nerve* 35:767–75

Smiley AK, Klingenberg JM, Boyce ST *et al.* (2006) Keratin expression in cultured skin substitutes suggests that the hyperproliferative phenotype observed *in vitro* is normalized after grafting. *Burns* 32:135–8

Uno H, Montagna W (1982) Reinnervation of hair follicle end organs and Meissner Corpuscles in skin grafts of Macaques. *J Invest Dermatol* 78:210–4

Woodbury CJ, Ritter AM, Koerber HR (2001) Central anatomy of individual rapidly adapting low-threshold mechanoreceptors innervating the 'hairy' skin of newborn mice: early maturation of hair follicle afferents. *J Comp Neurol* 436:304–23

Zhang Y, Andl T, Yang SH *et al.* (2008) Activation of beta-catenin signaling programs embryonic epidermis to hair follicle fate. *Development* 135:2161–72

# Filaggrin Null Mutations Are Not a Protective Factor for Acne Vulgaris

*Journal of Investigative Dermatology* (2011) **131,** 1378–1380; doi:10.1038/jid.2011.23; published online 17 February 2011

## TO THE EDITOR

Acne vulgaris is a very common skin disorder, affecting to some degree 88–94% of Singaporean adolescents (Tan *et al.*, 2007; Yosipovitch *et al.*, 2007). Genetic predisposition is a significant risk factor, as illustrated by familial and twin studies (Bataille *et al.*, 2002; Ghodsi *et al.*, 2009). The clinical features of acne include seborrhea, comedone formation, inflammatory pustules, nodules, and cysts, with resultant scarring. Important pathogenic mechanisms in acne include increased sebum production, hyperkeratinization and occlusion of the follicular duct, proliferation of *Propionibacterium acnes*, and an inflammatory reaction (Purdy and de Berker, 2006). *P. acnes* produces lipases, which liberate proinflammatory fatty acids from sebum and also triggers a cytokine response.

Filaggrin is expressed in terminally differentiating keratinocytes and has a key role in epithelial barrier formation. Immunostaining demonstrates increased filaggrin expression in the sebaceous duct and infundibulum of acne vulgaris skin (Kurokawa *et al.*, 1988), and *P. acnes* strains increase the expression of filaggrin and other differentiation-specific markers in normal human epidermal keratinocytes *in vitro* and in the suprabasal layers of human skin explants (Jarrousse *et al.*, 2007). Similarly, inflammatory cytokines resulted in increased filaggrin expression in sebaceous gland explants (Guy and Kealey, 1998). However, it is not known whether differences in filaggrin expression represent a primary or secondary effect in the pathogenesis of acne.

Null mutations in the filaggrin gene (*FLG*) result in reduced filaggrin expression and cause ichthyosis vulgaris (Smith *et al.*, 2006). Such mutations are common in the general population, being carried by ~10% of Europeans and 7.3% of Singaporean Chinese (Sandilands *et al.*, 2007; Chen *et al.*, personal communication). This high carrier rate in different populations suggests a heterozygote advantage, and it has been proposed that a more permeable skin barrier may have been beneficial in evolutionary history (Irvine and McLean, 2006). The co-existence of *FLG* null mutations with other gene mutations that disrupt epidermal differentiation may increase phenotype severity (Liao *et al.*, 2007; Gruber *et al.*, 2009). It is also possible that heterozygosity for null mutations has other effects on skin physiology. Studying a cohort of 284 European dermatology patients not selected for dry skin (Sergeant *et al.*, 2009) raised the possibility that carriage of one *FLG* null mutation could provide a protective effect against acne vulgaris. In the Sergeant

*Abbreviations:* FLG, *filaggrin gene*

**Table 1. Genotyping results and Fisher's exact test to investigate the association between *FLG* null mutations and acne vulgaris in a case–control study**

| *FLG* null mutation | Acne vulgaris cases and unselected population controls | Genotype[1] AA | Aa | aa | Total[2] | Fisher's exact test *P*-value |
|---|---|---|---|---|---|---|
| p.S406X | Cases | 279 | 0 | 0 | 279 | 0.524 |
| | Controls | 438 | 2 | 0 | 440 | |
| c.1249insG | Cases | 276 | 4 | 0 | 280 | 0.023 |
| | Controls | 440 | 0 | 0 | 440 | |
| c.2282del4 | Cases | 275 | 0 | 0 | 275 | 1.000 |
| | Controls | 436 | 1 | 0 | 437 | |
| c.3321delA | Cases | 277 | 4 | 0 | 281 | 0.582 |
| | Controls | 439 | 10 | 0 | 440 | |
| p.S1302X | Cases | 272 | 0 | 0 | 272 | 0.304 |
| | Controls | 434 | 4 | 0 | 438 | |
| p.S1515X | Cases | 275 | 1 | 0 | 276 | 1.000 |
| | Controls | 438 | 2 | 0 | 440 | |
| c.6834del5[3] | Cases | 280 | 1 | 0 | 281 | 0.390 |
| | Controls | 440 | 0 | 0 | 440 | |
| c.6950del8 | Cases | 278 | 3 | 0 | 281 | 1.000 |
| | Controls | 436 | 4 | 0 | 440 | |
| p.Q2417X | Cases | 278 | 2 | 0 | 280 | 0.151 |
| | Controls | 440 | 0 | 0 | 440 | |
| p.E2422X | Cases | 278 | 2 | 0 | 280 | 0.151 |
| | Controls | 440 | 0 | 0 | 440 | |
| c.7945delA | Cases | 274 | 2 | 0 | 276 | 0.562 |
| | Controls | 439 | 1 | 0 | 440 | |
| p.S2706X | Cases | 276 | 1 | 0 | 280 | 0.160 |
| | Controls | 431 | 7 | 0 | 438 | |
| c.8157delC[3] | Cases | 278 | 1 | 0 | 279 | 0.388 |
| | Controls | 440 | 0 | 0 | 440 | |
| p.R4307X | Cases | 278 | 1 | 0 | 279 | 0.488 |
| | Controls | 438 | 1 | 0 | 439 | |
| Combined *FLG* null genotype | Cases | 235 | 21 | 0 | 256 | 0.783 |
| | Controls | 402 | 31 | 1 | 434 | |

[1]AA, homozygous wild type for *FLG* null mutation; Aa, heterozygous for the stated *FLG* null mutation or, in the combined null genotype, any one of the screened mutations; aa, compound heterozygous (i.e., an individual having two different *FLG* null mutations). The rationale for generating this combined null genotype is based on the fact that each mutation results in premature termination of the profilaggrin molecule and hence absence of processed filaggrin (Sandilands *et al.*, 2007).
[2]The figures in the total column vary because of incomplete genotyping results; all available data have been used for optimal analysis of each individual variant, but the combined *FLG* null genotype data include only those individuals for whom all 14 genotype results are available.
[3]Previously unreported *FLG* mutations.

study, the odds ratio of acne in the carrier group was 0.3 (95% confidence interval 0.1–1.0), but the difference between these individuals and the group without *FLG* mutations did not reach statistical significance ($P=0.08$). In addition, this study relied on a recalled history of acne. We therefore aimed to test the hypothesis that *FLG* null mutations are protective against the development of acne vulgaris by studying a well-documented group of patients with acne, and comparing with a population control group.

A total of 287 Singaporean Chinese patients presenting with acne vulgaris to the National Skin Centre, a major dermatology outpatient facility in Singapore, were recruited: mean age 22 years (SD 4.8), range 14–50 (27% <20 years of age and 95% were under 30 years), 76.3%

were male. Acne vulgaris symptoms were reported for a mean of 5.6 years (SD 4.2), range from <1 to 32 years. Patients with polycystic ovarian syndrome were excluded from this study. Acne severity was assessed using the Global Severity Assessment Score (Lehmann *et al.*, 2002): 100 patients (34.8%) had mild acne, 129 (44.9%) had moderate acne, and 58 (20.3%) had severe acne. DNA samples from 440 unselected Singaporean Chinese population controls with a mean age of 44.6 years (SD 14.0), range 1–80, 44.1% male, for whom acne vulgaris status was unknown, were obtained from Singapore Bio-Bank, Singapore. This study was approved by the local domain specific ethical review board in accordance with the Declaration of Helsinki and all participants gave written, informed consent.

Cases and controls were screened for all 22 population-specific *FLG* null mutations as recently reported (Chen *et al.*, personal communication). In this group a total of 12 known *FLG* null mutations were detected: p.S406X, c.1249insG, c.2284del4, c.3321delA, p.S1302X, p.S1515X, c.6950del8, p.Q2417X, p.E2422X, c.7945delA, p.S2706X, and p.R4307X, plus two mutations: c.6834del5 and c.8157delC, which to our knowledge are previously unreported. Fisher's exact test and logistic regression analyses were used to compare the prevalence of *FLG* null mutations between cases and controls, using the statistical analysis package Stata (Version 9, Stata for Linux, StataCorp LP, College Station, TX). Power calculations were performed using Quanto version 1.2.4 (University of Southern California, http://hydra.usc.edu/gxe/).

In this study, 8.2% of Singaporean Chinese acne vulgaris cases carried one or more *FLG* null mutations compared with 7.3% of the control population, a non-significant difference (Fisher's exact test $P=0.783$, odds ratio 1.2, 95% confidence interval 0.7–2.1), shown in Table 1. It is unlikely that our failure to demonstrate an association has occurred because of lack of power, as analysis of 256 acne vulgaris patients and 434 controls, assuming a population prevalence of 88% for acne vulgaris of comparable severity (Tan *et al.*, 2007) would give a power of 99% to detect an odds ratio of 0.3 (Sergeant *et al.*, 2009) for the combined *FLG* null genotype, with two-sided $P=0.05$. The equivalent calculation using a population prevalence of 23% (based on Singaporean teenagers reporting treatment for their acne (Yosipovitch *et al.*, 2007)) gives an estimated power of 90% for this study. Furthermore, the comprehensive screening of all 22 reported *FLG* null mutations from this carefully characterized Singaporean Chinese population means that the lack of association is unlikely to have occurred because of incomplete ascertainment of the *FLG* genotype.

Our finding, that in this Singaporean Chinese population the frequency of *FLG* null mutations in acne vulgaris patients is not statistically different from the ethnically matched controls, indicates that filaggrin haploinsufficiency is unlikely to have a generic protective effect in acne. It is likely that the overexpression of filaggrin in the pilosebaceous units reported in the disorder is a bystander effect, reflecting other alterations in keratinocyte differentiation, but not itself critical for acne pathogenesis.

## CONFLICT OF INTEREST

The authors state no conflict of interest.

## ACKNOWLEDGMENTS

We would like to thank the affected individuals for their participation in this study. Work completed at the Institute of Medical Biology was funded by the Agency of Science, Technology and Research (A*STAR), S.J.B. is the recipient of a Wellcome Trust Intermediate Fellowship (ref 086398/Z/08/Z), and clinical research at the National Skin Centre is funded by an Enabling Grant from the National Medical Research Council, Singapore.

***John E.A. Common[1,5], Sara J. Brown[2,5], Rebecca L. Haines[1], Christabelle S.M. Goh[1], Huijia Chen[1,2], Anita Balakrishnan[1], Colin S. Munro[3], Audrey W.H. Tan[4], H.H. Tan[4], Mark B.Y. Tang[4] and E. Birgitte Lane[1]***

*[1]Epithelial Biology Group, Institute of Medical Biology, Immunos, Singapore; [2]Epithelial Genetics Group, Division of Molecular Medicine, Colleges of Life Sciences, Medicine, Dentistry and Nursing, Medical Sciences Institute, University of Dundee, Scotland, UK; [3]Alan Lyell Centre for Dermatology, Southern General Hospital, Glasgow, UK; [4]National Skin Centre, National Healthcare Group, Singapore*
*E-mail: birgit.lane@imb.a-star.edu.sg*
*[5]These authors contributed equally to this work.*

## REFERENCES

Bataille V, Snieder H, MacGregor AJ *et al.* (2002) The influence of genetics and environmental factors in the pathogenesis of acne: a twin study of acne in women. *J Invest Dermatol* 119:1317–22

Ghodsi SZ, Orawa H, Zouboulis CC (2009) Prevalence, severity, and severity risk factors of acne in high school pupils: a community-based study. *J Invest Dermatol* 129:2136–41

Gruber R, Wilson NJ, Smith FJD *et al.* (2009) Increased pachyonychia congenita severity in patients with concurrent keratin and filaggrin mutations. *Br J Dermatol* 161:1391–5

Guy R, Kealey T (1998) The effects of inflammatory cytokines on the isolated human sebaceous infundibulum. *J Invest Dermatol* 110: 410–5

Irvine AD, McLean WH (2006) Breaking the (un)sound barrier: filaggrin is a major gene for atopic dermatitis. *J Invest Dermatol* 126:1200–2

Jarrousse V, Castex-Rizzi N, Khammari A *et al.* (2007) Modulation of integrins and filaggrin expression by Propionibacterium acnes extracts on keratinocytes. *Arch Dermatol Res* 299:441–7

Kurokawa I, Mayer-da-Silva A, Gollnick H *et al.* (1988) Monoclonal antibody labeling for cytokeratins and filaggrin in the human pilosebaceous unit of normal, seborrhoeic and acne skin. *J Invest Dermatol* 91:566–71

Lehmann HP, Robinson KA, Andrews JS *et al.* (2002) Acne therapy: a methodologic review. *J Am Acad Dermatol* 47:231–40

Liao H, Waters AJ, Goudie DR *et al.* (2007) Filaggrin mutations are genetic modifying factors exacerbating X-linked ichthyosis. *J Invest Dermatol* 127:2795–8

Purdy S, de Berker D (2006) Acne. *BMJ* 333: 949–53

Sandilands A, Terron-Kwiatkowski A, Hull PR *et al.* (2007) Comprehensive analysis of the gene encoding filaggrin uncovers prevalent and rare mutations in ichthyosis vulgaris and atopic eczema. *Nat Genet* 39:650–4

Sergeant A, Campbell LE, Hull PR *et al.* (2009) Heterozygous null alleles in filaggrin contribute to clinical dry skin in young adults and the elderly. *J Invest Dermatol* 129:1042–5

Smith FJD, Irvine AD, Terron-Kwiatkowski A *et al.* (2006) Loss-of-function mutations in the gene encoding filaggrin cause ichthyosis vulgaris. *Nat Genet* 38:337–42

Tan HH, Tan AW, Barkham T *et al.* (2007) Community-based study of acne vulgaris in adolescents in Singapore. *Br J Dermatol* 157:547–51

Yosipovitch G, Tang M, Dawn AG *et al.* (2007) Study of psychological stress, sebum production and acne vulgaris in adolescents. *Acta Derm Venereol* 87:135–9

ORIGINAL ARTICLE SKIN AND EYE DISEASES

# Levels of filaggrin degradation products are influenced by both filaggrin genotype and atopic dermatitis severity

S. Kezic*[,1], G. M. O'Regan*[,2,3], N. Yau[1], A. Sandilands[4], H. Chen[4,5], L. E. Campbell[4], K. Kroboth[4], R. Watson[2], M. Rowland[2,6], W. H. Irwin McLean[4] & A. D. Irvine[2,3,7]

[1]Coronel Institute of Occupational Health, Academic Medical Center, Amsterdam, the Netherlands; [2]The National Children's Research Centre, Our Lady's Children's Hospital; [3]Department of Paediatric Dermatology, Our Lady's Children's Hospital, Dublin, Ireland; [4]Division of Molecular Medicine, University of Dundee, Dundee, UK; [5]Institute of Medical Biology, Agency for Science, Technology and Research, Singapore; [6]UCD School of Medicine and Medical Sciences, Dublin, Ireland; [7]Department of Medicine, Trinity College, Dublin, Ireland

**To cite this article:** Kezic S, O'Regan GM, Yau N, Sandilands A, Chen H, Campbell LE, Kroboth K, Watson R, Rowland M, Irwin McLean WH, Irvine AD. Levels of filaggrin degradation products are influenced by both filaggrin genotype and atopic dermatitis severity. *Allergy* 2011; **66**: 934–940.

**Keywords**
2-pyrrolidone-5-carboxylic acid; FLG gene mutations; natural moisturizing factor; tyrosine; urocanic acid.

**Correspondence**
Alan Irvine, Department of Medicine, Trinity College, Dublin, Ireland.
Tel.: +3531 428 2532
Fax: +3531 428 2651
E-mail: irvinea@tcd.ie
or:
Dr S. Kezic, Academic Medical Center, Coronel Institute, room K0-104, PO Box 22 700, the Netherlands, 1100 DE Amsterdam.
Tel.: +31 20 5665321
Fax: +31 20 6977161
E-mail: s.kezic@amc.uva.nl

*Contributed equally

Accepted for publication 12 December 2010

DOI:10.1111/j.1398-9995.2010.02540.x

Following standard genetic practice, in this paper, *FLG−/−* designates a patient homozygous for null alleles (i.e. two null alleles); *FLG+/−* a heterozygote null allele/wild type (i.e. one null allele) and *FLG +/+* a homozygote wild type (i.e. no null alleles). A further abbreviation describes AD patients with *FLG* mutations (*FLG+/−* and *FLG−/−*) as $AD_{FLG}$ and those without *FLG* mutations (i.e. *FLG+/+*) as $AD_{NON\text{-}FLG}$.

Edited by: Thomas Bieber

## Abstract

**Background:** Filaggrin, coded by *FLG*, is the main source of several major components of natural moisturizing factor (NMF) in the stratum corneum (SC), including pyrrolidone carboxylic acid (PCA) and urocanic acid (UCA). Loss-offunction mutations in *FLG* lead to reduced levels of filaggrin degradation products in the SC. It has recently been suggested that expression of filaggrin may additionally be influenced by the atopic inflammatory response. In this study, we investigated the levels of several breakdown products of filaggrin in the SC in healthy controls (CTRL) and patients with atopic dermatitis (AD) in relation to *FLG* null allele status. We examined the relationship between NMF (defined here as the sum of PCA and UCA) and AD severity.

**Methods:** The SC levels of filaggrin degradation products including PCA, UCA, histidine (HIS) and tyrosine were determined in 24 CTRL and 96 patients with moderate-to-severe AD. All subjects were screened for 11 *FLG* mutations relevant for the study population.

**Results:** The levels of PCA, UCA and HIS correlated with *FLG* genotype. Furthermore, these levels were higher in the CTRL when compared to AD patients with no *FLG* mutations. Multiple regression analysis showed that NMF levels were independently associated with *FLG* genotype and severity of disease.

**Conclusion:** Decreased NMF is a global feature of moderate-to-severe AD; within AD, *FLG* genotype is the major determinant of NMF, with disease severity as a secondary modifier. NMF components are reliably determined by a noninvasive and relatively inexpensive tape stripping technique.

**Abbreviations**
*FLG*, Filaggrin gene; HIS, histidine; NMF, natural moisturizing factor (defined in this study as the sum of PCA and UCA); PCA, pyrrolidone carboxylic acid; SC, stratum corneum; TEWL, transepidermal water loss; TYR, tyrosine; UCA, urocanic acid.

Filaggrin gene (*FLG*) loss-of-function mutations underlie ichthyosis vulgaris, a semi-dominant disorder of keratinization, and are the strongest and most widely replicated genetic risk factor for atopic dermatitis (AD) (1). We have recently shown

using Raman spectroscopy that *FLG* genotype is a major determinant of natural moisturizing factor (NMF) in the stratum corneum (SC) (2, 3). Furthermore, we found that the tyrosine (TYR) levels in the palmar SC of patients with AD were positively correlated with the number of *FLG* null alleles, although the mechanism underlying elevated TYR in the carriers of *FLG* loss-of-function mutations is not yet clear.

*FLG* codes for profilaggrin, a large protein precursor comprised of 10–12 filaggrin repeats. The most abundant amino acid residues in filaggrin repeats are basic amino acids histidine (HIS) (413/4061 residues; 10.17%) and arginine (440/4061; 10.83%) and the polar residue glutamine (367/4061; 9.04%) (Fig. S1). The protein is also significantly rich in glycine (12.76%) and serine (24.06%), consistent with the structural similarity of filaggrin and related proteins such as loricrin to the variable domains of keratins, which largely consist of glycine/serine loop structures (4). Filaggrin contains fewer than average TYR residues (1.28%); however, the linker segments and carboxyl-terminal domains of profilaggrin contain dense and highly conserved TYR-rich motifs (5–7). In the later stages of epidermal differentiation, filaggrin is degraded into free amino acids and their derivatives; a major proportion of the total SC free amino acids (70–100%) are derived from filaggrin (8–10). Glutamine is further converted via a nonenzymatic process into pyrrolidone-5-carboxylic acid (PCA). PCA is highly hygroscopic and is one of the major NMF constituents (11). Histidine is deiminated to trans-urocanic acid (trans-UCA) by the catalytic action of histidase (12). Trans-UCA, which is converted to cis-UCA upon UV irradiation, functions as a major chromophore in skin and exhibits immunomodulatory effects (13). Furthermore, UCA has been suggested as a contributor to the maintenance of the acidic pH in skin, which is crucial for the activity of several key enzymes within the SC (14). Recently, we have shown that UCA and PCA exert profound effects on *Staphylococcus aureus* at physiologic concentrations (15). Filaggrin degradation products thus have multiple functions, and reduced levels as a consequence of *FLG* loss-of-function mutations may affect skin hydration, local immune responses, lipid composition (16, 17) and maintenance of epidermal homeostasis.

*In vitro* evidence has suggested that co-expression of IL-4 and IL-13, known to be upregulated in AD, can downregulate the expression of filaggrin (18). Here, we compared the levels of filaggrin degradation products in moderate-to-severe AD patients with known *FLG* mutations ($AD_{FLG}$) and those wild type for *FLG* mutations ($AD_{NON\text{-}FLG}$) with unaffected controls. We also investigated the relationship between severity of AD and NMF levels, defined here as the sum of PCA and UCA.

## Subjects and methods

Ninety-six unrelated Irish children with a history of moderate-to-severe AD were recruited from dedicated secondary and tertiary referral AD clinics. Twenty-five healthy children with no history of AD or atopic disease were recruited from patients attending for laser therapy of vascular birthmarks. The diagnosis and phenotyping of AD was made by experienced pediatric dermatologists. All subjects met the UK diagnostic criteria (19). Exclusion criteria from the study were patients who had received systemic therapy, such as corticosteroids or immunosuppressants in the preceeding 3 months, and patients whose ancestry was not exclusively Irish (4/4 grandparents). Detailed phenotypic data were collected and are presented in Table 1. The Nottingham Eczema Severity Score (NESS) was selected as a measure of disease severity (20). The study was conducted in accordance with Helsinki Declarations and was approved by the Research Ethics Committee of Our Lady's Children's Hospital, Dublin. Full written consent was obtained from all patients or their parents. Transepidermal

**Table 1** Clinical and demographic characteristics of the cohort

| *FLG* genotype | Number | Age in years *n* mean (SD) | Male gender *n* (%) | IgE *n* median (range) | NESS *n* median (range) | SCORAD* *n*, mean median (range) | TEWL *n* median (range) ($g/m^2/h$) |
|---|---|---|---|---|---|---|---|
| CTRL | 24 | 24<br>5.6 (4.0) | 9 (38) | / | / | / | 24<br>10.4 (6.8–10.7) |
| *FLG+/+* | 40 | 34<br>8.1 (4.2) | 19 (53) | 29<br>599<br>(2–44294) | 28<br>13 (8–15) | 9<br>12.0 (0–36) | 26<br>12.0 (5.2–38.3) |
| *FLG+/–* | 37 | 37<br>8.7 (4.1) | 23 (62) | 35<br>1033<br>(14.9–23260) | 36<br>12 (6–15) | 19<br>12 (0–28) | 35<br>14.1 (5.3–30.6) |
| *FLG–/–* | 19 | 19<br>8.6 (4.5) | 12 (63) | 16<br>1803<br>(62.9–6684) | 18<br>12 (6–15) | 10<br>21 (0–32) | 17<br>15.3 (8.5–30.4) |
| Total *AD* | 96 | 90<br>8.5 (4.2) | 54 (59.0) | 80<br>1016 (2–44294) | 82<br>12 (6–15) | 38<br>13 (0–36) | 78<br>13.5 (5.2–38.3) |

AD, atopic dermatitis; CTRL, healthy controls; TEWL, transepidermal water loss; NESS, Nottingham Eczema Severity Score (20).
*Not available for all subjects.

water loss (TEWL) was measured on nonlesional skin of the extensor forearm (Tewameter 300; Courage and Khazaka Electronic GmbH, Cologne, Germany).

### Genetic screening

All patients were screened for 11 *FLG* mutations found in the Irish population, as previously described (3). Based on screening for these 11 prevalent mutations, 58% were carriers of one or more *FLG* mutations (42% *FLG*+/+; 38% *FLG*+/−; 20% *FLG*−/−). Primers and conditions for these 11 mutations are given in detail in the Supporting information.

### Determination of filaggrin degradation products in the stratum corneum

The levels of PCA, trans- and cis-UCA, HIS and TYR were measured using a tape stripping technique from nonlesional skin of the volar forearm as previously described by Kezic et al. (21). Briefly, round adhesive tape discs (3.8 $cm^2$, D-Squame; CuDerm, Dallas, TX, USA) were attached to the skin of the forearm. Each tape was applied to the volar aspect of the forearm for 10 s using a disc pressure applicator (CuDerm). The tape strip was gently removed with tweezers and stored in a closed vial at −20°C until analysis. The first strip was discarded as it may have contained dirt and remnants of cosmetic products; the second, third, fourth and fifth tape strips were applied on the same skin spot. The collected four tape strips were cut into two equal pieces. For the analysis, halves of four strips were pooled for the analysis. Before HPLC analysis, 500 μl KOH solution, 0.1 M, was added to the tape strips, followed by 2 h of continuous shaking. The alkaline extracts were neutralized with 3 μl perchloric acid, 12 M, and filtered through a 0.2-μm membrane filter. An aliquot of 50 μl was taken for the analysis of proteins by using the Bio-Rad DC protein microassay (Bio-Rad Laboratories, München, Germany) using commercially available BSA for standardization. Another 50 μl of aliquot was introduced into the HPLC system containing a 250 × 3 mm reversed-phase Prevail HPLC column (Grace/Alltech, Breda, the Netherlands). For the analysis of PCA and UCA isomers, isocratic elution was performed with 20 mM ammonium formate containing 1.5 mM tetrabutylammonium hydroxide and 1% acetonitrile at pH 7.3 (HPLC method I). The effluent was monitored at 210 nm for PCA and at 270 nm for both UCA isomers, using two sequential UV/Vis detectors, model 759A (Applied Biosystems, Carlsbad, California 92008) and model UV-975 (Jasco, Tokyo, Japan). The levels of the UCA were given as a total of the cis- and trans-isomer. Representative chromatograms obtained from the SC of patients of different *FLG* status are shown in Fig. S2. For the analysis of HIS and TYR, the HPLC analysis was modified as follows: isocratic elution was performed with 6 mM hydrochloric acid, 0.3 mM octane-1-sulfonic acid sodium salt and 1% acetonitrile at pH 3 (HPLC method II). HIS and TYR were monitored at 210 nm using a UV/VIS-detector (Jasco, model UV975). The injection volume was 4 μl. Representative chromatograms obtained from the SC of patients of different *FLG* status are shown in Fig. S3. The levels of PCA, UCA, HIS and TYR were normalized for the protein value.

### Statistical analysis

Patients were characterized, *a priori*, into three genotypes (*FLG*+/+, *FLG*+/− and *FLG*−/−), where *FLG*+/+ represents patients with no *FLG* mutations, *FLG*+/− represents patients with one *FLG* mutation, and *FLG*−/− represents patients with two *FLG* mutations. The healthy controls (CTRL) were genotyped in the same way as patients with AD; all but one control individual were wild type for *FLG* mutations and this person was excluded from the data analysis.

For testing of distribution, we used the Shapiro–Wilk's test. In case of non-normal distribution, data were log-transformed. The means were then compared by analysis of variance (ANOVA) followed by a *post hoc* Tukey analysis for comparison of pairwise differences between the genotype subgroups of patients with AD. For comparison of the levels of filaggin degradation products between CTRL *vs* *FLG*+/+, we used a Student's one-tailed test. *P*-values $< 0.05$ were considered statistically significant. In the figures, log-normal data are presented as geometric mean ± 95% CIs. To investigate the contribution of the number of *FLG* null mutations and severity of disease, we used multiple regression analysis with NMF (the sum of PCA and UCA) as a dependent variable and *FLG* mutations (0, 1 and 2) and disease severity as assessed by NESS and the SCORing Atopic Dermatitis (SCORAD) score as independent variables (SPSS 7.0, IBM Corporation, Somers, NY 10589).

**Table 2** Median values (and range) of filaggrin degradation products in healthy controls (CTRL) and patients with atopic dermatitis in relation to *FLG* genotype

| | PCA (mmol/g protein) | UCA (mmol/g protein) | HIS (mmol/g protein) | NMF (PCA + UCA) (mmol/g protein) | PCA + UCA + HIS (mmol/g protein) |
|---|---|---|---|---|---|
| *N* | 120 | 120 | 97 | 120 | 97 |
| CTRL | 1.20 (0.43–3.83) | 0.26 (0.10–1.29) | 0.12 (0.02–0.23) | 1.60 (0.54–4.8) | 1.64 (0.74–4.64) |
| *FLG*+/+ | 0.85 (0.29–3.62) | 0.20 (0.04–0.77) | 0.08 (0.007–0.21) | 1.09 (0.34–4.10) | 1.20 (0.40–3.60) |
| *FLG*+/− | 0.40 (0.09–2.07) | 0.15 (0.03–0.45) | 0.05 (0.001–0.25) | 0.84 (0.19–2.34) | 0.80 (0.20–1.90) |
| *FLG*−/− | 0.23 (0.09–0.94) | 0.07 (0.008–0.32) | 0.01 (0.002–0.18) | 0.36 (0.12–1.12) | 0.40 (0.10–1.30) |

HIS, histidine; NMF, natural moisturizing factor; PCA, pyrrolidone carboxylic acid; UCA, urocanic acid.

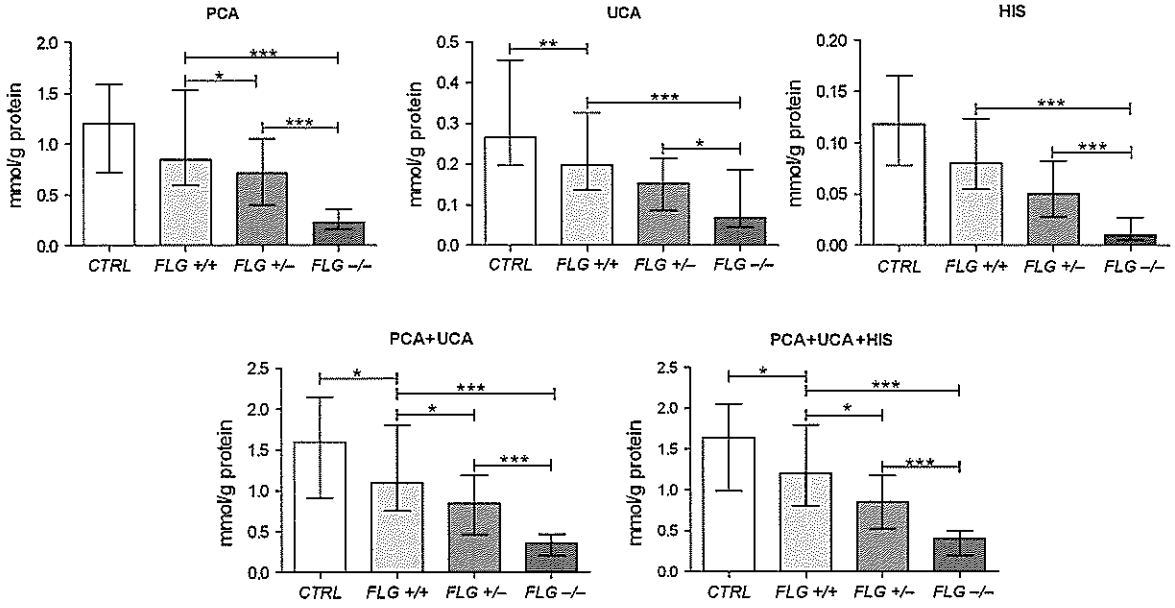



**Figure 1** Filaggrin degradation products (geometric mean and 95% CI) in healthy controls (CTRL) and patients with atopic dermatitis (AD) in relation to *FLG* genotype. For statistic analysis, all data were log-transformed. Within AD patient subgroups, anova analysis with a *post hoc* Tukey multiple correction has been applied. To test differences between CTRL and $AD_{NON\text{-}FLG}$ (*FLG*+/+), the one-tailed Student's *t*-test was used. $^{*}P < 0.05$, $^{**}P < 0.01$; $^{***}P < 0.001$.

## Results

### Filaggrin degradation components in relation to *FLG* allele status

Clinical characteristics and summary data of the study cohort including *FLG* genotype are outlined in Table 1. The median values of PCA, UCA, HIS and their sum according to *FLG* genotype are shown in Table 2. Within subgroups, a range of values of filaggrin breakdown products was seen (Table 2); these values showed log-normal distribution. *Post hoc* Tukey analysis showed clearly that the levels of these filaggrin breakdown products in patients with AD were related to *FLG* status (Fig. 1).

As seen from Table 1, the average age and proportion of men in patients with AD were higher than in the controls. However, in the multiple regression analysis with *FLG* status, age and sex as independent variables, sex and age did not have an effect on the NMF ($P > 0.1$).

The concentration of filaggrin degradation products is normalized by the protein amount to compensate for variable amounts of SC harvested by tape stripping. To enable comparison with literature data, we report the mean concentrations of PCA and UCA in nmol/cm$^2$. The PCA levels amounted to 17.9, 14.4, 8.2 and 4.7 nmol/cm$^2$ for CTRL, *FLG*+/+, *FLG*+/– and *FLG*–/–, respectively. The respective UCA levels were 4.2, 2.4, 1.8 and 1.4 nmol/cm$^2$. Approximating the thickness of the SC as 10 μm, the values expressed in nmol/cm$^2$ correspond to those in mM (21). The concentration of UCA found in CTRL (4.2 nmol/cm$^2$, i.e. 4.2 mM) is in good agreement with published results (22).

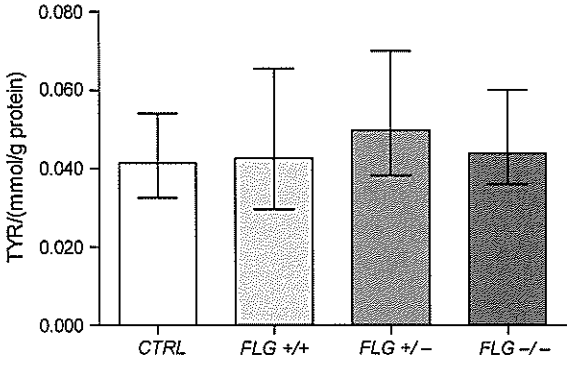



**Figure 2** Levels of tyrosine (geometric mean and 95% CI) in relation to *FLG* genotype.

In contrast to PCA, UCA and HIS, the levels of TYR did not correlate with the *FLG* null allele status ($P > 0.1$, *post hoc* Tukey) (Fig. 2). A significant association between the levels of PCA + UCA + HIS and TYR was obtained in CTRL and *FLG* +/+ AD patients (Fig. 3A). However, the correlation between UCA + PCA + HIS and TYR was less striking in the *FLG*+/– and *FLG*–/– patients (Fig. 3B). Also, the slope of the regression line was less steep in $AD_{FLG}$ patients, i.e. the level of TYR for any given concentration of

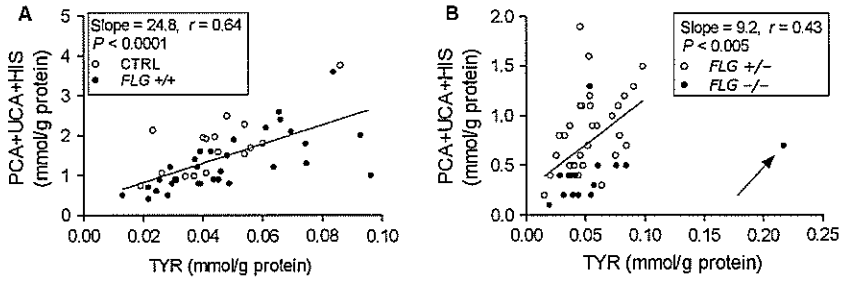



**Figure 3** Correlation between the levels of tyrosine and pyrrolidone carboxylic acid + urocanic acid + histidine in (A) healthy controls and *FLG+/+* and in (B) *FLG+/−* and *FLG−/−*. The outlier in (B) was excluded in the linear regression analysis. *r*, Pearson's correlation coefficient.

PCA + UCA + HIS was higher in $AD_{FLG}$ than in the CTRL or $AD_{NON\text{-}FLG}$ subjects.

### Natural moisturizing factor levels in relation to individual *FLG* mutations

Among the 11 investigated mutations, R501X and 2282del4 mutations were the most common. In the *FLG+/−* subgroup, there were 14 carriers of R501X and 15 carriers of 2282del4. Both mutations led to similar mean levels of PCA+UCA (0.87 ± 0.39 and 1.09 ± 0.62, respectively; two-sided Student's *t*-test, $P > 0.05$). For other mutations, we had only a limited number of carriers and detailed mutation-specific expression comparisons were not possible.

### Influence of presence of atopic dermatitis and disease severity on natural moisturizing factor levels

To investigate whether the levels of filaggrin degradation products were influenced, in addition to *FLG* genotype, by the presence of disease *per se*, we compared these levels between CTRL and *FLG* wild-type patients with AD ($AD_{NON\text{-}FLG}$; *FLG+/+*). As shown in Fig. 1 and Table 1, the levels of all NMF components tended to be higher in the CTRL subgroup when compared to the $AD_{NON\text{-}FLG}$ (*FLG+/+*) subgroup: for UCA, PCA + UCA and PCA + UCA + HIS; this difference was statistically significant. To investigate the individual contribution of *FLG* null mutations and disease severity on the NMF levels, we performed multiple regression analysis in AD patients with NMF as the dependent variable and *FLG* mutations (0, 1 and 2) and disease severity as assessed either by NESS ($n = 82$) or by SCORAD ($n = 38$) as independent variables. The results showed that in addition to *FLG* mutations, severity of disease significantly contributes toward NMF levels (Table 3). There was no significant association between NMF components and TEWL or IgE in any of AD subgroups. Furthermore, the disease severity and the levels of IgE and TEWL did not statistically differ between the three AD subgroups ($P > 0.1$, ANOVA and Tukey's multiple comparison test). Predictably, patients with AD showed elevated TEWL when compared to CTRL (one-sided Student's *t*-test of log-transformed data; $P < 0.05$); however, there was no difference in TEWL between AD subgroups.

**Table 3** Natural moisturizing factor: relationship with *FLG* mutations and disease severity as assessed by NESS and SCORAD (multiple linear regressions)

| | Unstandardized coefficients B (95% CI) | Standardized coefficient beta | Significance, *P* |
|---|---|---|---|
| Constant | 2.248 (1.527; −2.968) | | 0.000 |
| *FLG* MUT | −0.549 (−0.729; −0.369) | −0.556 | 0.000 |
| NESS | −0.068 (−0.124–0.011) | −0.218 | 0.010 |
| Constant | 1.301 (1.02; 71.575) | | 0.000 |
| *FLG* MUT | −0.333 (0.504; −0.162) | −0.512 | 0.000 |
| SCORAD | −0.016 (−0.030; −0.002 | −0.309 | 0.010 |

NESS, Nottingham Eczema Severity Score (19); *FLG* MUT, number of *FLG* mutations (0, 1 and 2).

## Discussion

Since the discovery of the strong association between *FLG* mutations and AD, the mechanisms through which these mutations lead to disease causation and modification have been of great interest. We have previously shown that *FLG* genotypes are the major determinants of NMF and that this association is sufficiently strong to define three SC endophenotypes within AD (3). In this study, we have replicated these findings using a SC tape stripping technique followed by HPLC as an alternative noninvasive method. We then examined the relationship between AD patients with no *FLG* mutations ($AD_{NON\text{-}FLG}$) and normal controls. By very carefully defining the *FLG* status of a well-characterized collection of children with moderate-to-severe AD, we were able to further examine the relationship between AD severity, *FLG* genotype and NMF expression. We found that the two commonest *FLG* mutations had similar effects on NMF, consistent with our previous protein expression data (23) demonstrating that these mutations appear to have biologically equivalent effects. Strongly reduced levels of the filaggrin degradation products

in *FLG−/−* patients (threefold decrease in *FLG−/−* patients when compared to *FLG+/+* patients) imply that filaggrin is the major source of HIS, PCA and UCA in the SC, which is consistent with previous work (10). Thus, *FLG* genotype is the major determinant of filaggrin-derived NMF components.

Here, we demonstrated that, even in the absence of *FLG* mutations, NMF is significantly reduced in the nonlesional skin of patients with AD, implying a systemic downregulation of *FLG* expression. This is consistent with the *in vitro* findings of Howell et al. (18) and highlights the complex interplay between the skin barrier and a systemic immunologic process. Consistent with these findings, we also showed that, in AD addition to *FLG* genotype status, AD severity also has an effect on NMF levels. These findings are important and may have clinical relevance as they suggest decreased NMF is a general finding in AD and that upregulation of *FLG* expression may be beneficial to all patients with AD, rather than having an effect limited to those with *FLG* mutations. Measurement of NMF may be of value as an investigational tool to measure response to treatment before and after treatment of AD with alternative treatment strategies. HPLC of tape strip–derived material is an efficient way to measure the levels of multiple NMF components in such studies.

In contrast to our recent study reporting a markedly high positive correlation of TYR with the number of *FLG* null alleles (3), this present study observes similar levels of TYR in all investigated groups. There are several methodological differences that might have caused discrepancy in the results between these two studies. Our previous study was based on Raman spectroscopy, while here we determined TYR by using HPLC and UV spectroscopy. Both methods are highly specific, and observed differences are not likely to be caused by the use of these different detection methods. However, in the present study, TYR levels were measured in the uppermost layers of the SC from the forearm skin while in the Raman study, measurements were taken on the palmar skin and the TYR levels were averaged for the entire depth of the SC. The strong relationship between filaggrin degradation products and TYR in CTRL and *FLG+/+* found in the present study suggests that TYR in the SC may originate from filaggrin, although filaggrin is not particularly TYR rich. In $AD_{FLG}$, this relationship is less striking. Furthermore, the ratio of TYR to filaggrin degradation products is higher in $AD_{FLG}$ when compared to CTRL and $AD_{NON\text{-}FLG}$, suggesting either accumulation of TYR in the SC or alternative pathways in AD patients with *FLG* mutations which leads to elevated TYR levels. Further study on these TYR levels and other filaggrin degradation products in relation to skin location and SC depth is needed to clarify the elevated levels of TYR in the carriers of *FLG* loss-of-function mutations. It is quite possible that the Raman TYR signal observed in *FLG* mutation carriers, which is particularly in homozygotes (3), is because of upregulation and/or unmasking of another TYR-rich protein as part of a compensatory response to *FLG* deficiency. Proteomics analysis of SC from individuals with defined *FLG* genotypes may help resolve this issue in the future.

## Supporting Information

Additional Supporting Information may be found in the online version of this article at www.wileyonlinelibrary.com

**Data S1.** FLG genotyping.
**Figure S1.** Amino acid composition of human profilaggrin.
**Figure S2.** The representative HPLC chromatograms (HPLC method I) of the extracts of the SC of patients of different *FLG* status and of the standard solution of PCA, trans- and cis-UCA (25, 25 and 75 μM, respectively).
**Figure S3.** The representative HPLC chromatograms (HPLC method II) of the extracts of the SC of patients of different *FLG* status and of the standard solution of HIS and TYR (20 and 25 μM, respectively).

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.

## References


1. Sandilands A, Sutherland C, Irvine AD, McLean WH. Filaggrin in the frontline: role in skin barrier function and disease. *J Cell Sci* 2009;**122**:1285–1294.
2. Kezic S, Kemperman PM, Koster ES, de Jongh CM, Thio HB, Campbell LE et al. Loss-of-function mutations in the filaggrin gene lead to reduced level of natural moisturizing factor in the stratum corneum. *J Invest Dermatol* 2008;**128**:2117–2119.
3. O'Regan GM, Kemperman PM, Sandilands A, Chen H, Campbell LE, Kroboth K et al. Raman profiles of the stratum corneum define 3 filaggrin genotype-determined atopic dermatitis endophenotypes. *J Allergy Clin Immunol* 2010;**126**:574–580.
4. Korge BP, Gan SQ, McBride OW, Mischke D, Steinert PM. Extensive size polymorphism of the human keratin 10 chain resides in the C-terminal V2 subdomain due to variable numbers and sizes of glycine loops. *Proc Natl Acad Sci USA* 1992;**89**:910–914.
5. Gan SQ, McBride OW, Idler WW, Markova N, Steinert PM. Organization, structure, and polymorphisms of the human profilaggrin gene. *Biochemistry* 1990;**29**:9432–9440.
6. Presland RB, Haydock PV, Fleckman P, Nirunsuksiri W, Dale BA. Characterization of the human epidermal profilaggrin gene. Genomic organization and identification of an S-100-like calcium binding domain at the amino terminus. *J Biol Chem* 1992;**267**:23772–23781.
7. Resing KA, Walsh KA, Haugen-Scofield J, Dale BA. Identification of proteolytic cleavage sites in the conversion of profilaggrin to filaggrin in mammalian epidermis. *J Biol Chem* 1989;**264**:1837–1845.
8. Scott IR, Harding CR, Barrett JG. Histidine-rich protein of the keratohyalin granules. Source of the free amino acids, urocanic acid and pyrrolidone carboxylic acid in the stratum corneum. *Biochim Biophys Acta* 1982;**719**:110–117.
9. Scott IR, Harding CR. Filaggrin breakdown to water binding compounds during development of the rat stratum corneum is controlled by the water activity of the environment. *Dev Biol* 1986;**115**:84–92.

10. Rawlings AV, Scott IR, Harding CR, Bowser PA. Stratum corneum moisturization at the molecular level. *J Invest Dermatol* 1994; **103**:731–741.
11. Barrett JG, Scott IR. Pyrrolidone carboxylic acid synthesis in guinea pig epidermis. *J Invest Dermatol* 1983;**81**:122–124.
12. Baden HP, Pathak MA. The metabolism and function of urocanic acid in skin. *J Invest Dermatol* 1967;**48**:11–17.
13. Gibbs NK, Tye J, Norval M. Recent advances in urocanic acid photochemistry, photobiology and photoimmunology. *Photochem Photobiol Sci* 2008;7:655–667.
14. Elias PM, Steinhoff M. "Outside-to-inside" (and now back to "outside") pathogenic mechanisms in atopic dermatitis. *J Invest Dermatol* 2008;**128**:1067–1070.
15. Miajlovic H, Fallon PG, Irvine AD, Foster TJ. Effect of filaggrin breakdown products on growth of and protein expression by *Staphylococcus aureus*. *J Allergy Clin Immunol* 2010;**126**:1184–90.e3.
16. Jungersted JM, Scheer H, Mempel M, Baurecht H, Cifuentes L, Høgh JK et al. Stratum corneum lipids, skin barrier function and filaggrin mutations in patients with atopic eczema. *Allergy* 2010;65:911–918.
17. Sergeant A, Campbell LE, Hull PR, Porter M, Palmer CNA, Smith FJD et al. Heterozygous null alleles in filaggrin contribute to clinical dry skin in young adults and the elderly. *J Invest Dermatol* 2009;**129**:1042–1045.
18. Howell MD, Kim BE, Gao P, Grant AV, Boguniewicz M, Debenedetto A et al. Cytokine modulation of atopic dermatitis filaggrin skin expression. *J Allergy Clin Immunol* 2007;**120**:150–155.
19. Williams HC, Burney PG, Hay RJ, Archer CB, Shipley MJ, Hunter JJ et al. The U.K. Working Party's Diagnostic Criteria for Atopic Dermatitis. I. Derivation of a minimum set of discriminators for atopic dermatitis. *Br J Dermatol* 1994;**131**: 383–396.
20. Emerson RM, Charman CR, Williams HC. The Nottingham Eczema Severity Score: preliminary refinement of the Rajka and Langeland grading. *Br J Dermatol* 2000; **142**:288–297.
21. Kezic S, Kammeyer A, Calkoen F, Fluhr JW, Bos JD. Natural moisturizing factor components in the stratum corneum as biomarkers of filaggrin genotype: evaluation of minimally invasive methods. *Br J Dermatol* 2009;**161**: 1098–1104.
22. Safer D, Brenes M, Dunipace S, Schad G. Urocanic acid is a major chemoattractant for the skin-penetrating parasitic nematode Strongyloides stercoralis. *Proc Natl Acad Sci USA* 2007;**104**:1627–1630.
23. Sandilands A, Terron-Kwiatkowski A, Hull PR, O'Regan GM, Clayton TH, Watson RM et al. Comprehensive analysis of the gene encoding filaggrin uncovers prevalent and rare mutations in ichthyosis vulgaris and atopic eczema. *Nat Genet* 2007;**39**: 650–654.

DOI:10.1111/j.1600-0625.2010.01109.x
www.blackwellpublishing.com/EXD

Letter to the Editor

# Missing C-terminal filaggrin expression, NFkappaB activation and hyperproliferation identify the dog as a putative model to study epidermal dysfunction in atopic dermatitis

**Ludovic Chervet[1,4], Arnaud Galichet[1,4], W. H. Irwin McLean[2], Huijia Chen[2], Maja M. Suter[1,4], Petra J. Roosje[3,4*] and Eliane J. Müller[1,4*]**

[1]Molecular Dermatology, Institute of Animal Pathology, Vetsuisse Faculty, University of Bern, Bern, Switzerland;
[2]Epithelial Genetics Group, Division of Molecular Medicine, Colleges of Life Sciences and Medicine, Dentistry & Nursing, University of Dundee, Dundee, Scotland;
[3]Division of Clinical Dermatology, Department of Clinical Veterinary Medicine, University of Bern, Bern, Switzerland;
[4]DermFocus, Vetsuisse Faculty, University of Bern, Bern, Switzerland
*Correspondence*: Eliane J. Müller, Molecular Dermatology, Institute of Animal Pathology, Vetsuisse Faculty, University of Bern, Bern, Switzerland, Länggass-Strasse 122, Postfach 8466, Bern CH-3001, Switzerland, Tel.: +41 31 631 24 03, Fax: +41 31 631 26 35, e-mail: eliane.mueller@itpa.unibe.ch
*Equal contribution.

**Abstract:** Filaggrin loss-of-function mutations resulting in C-terminal protein truncations are strong predisposing factors in human atopic dermatitis (AD). To assess the possibility of similar truncations in canine AD, an exclusion strategy was designed on 16 control and 18 AD dogs of various breeds. Comparative immunofluorescence microscopy was performed with an antibody raised against the canine filaggrin C-terminus and a commercial N-terminal antibody. Concurrent with human AD-like features such as generalized NFKB activation and hyperproliferation, four distinctive filaggrin expression patterns were indentified in non-lesional skin. It was found that 10/18 AD dogs exhibited an identical pattern for both antibodies with comparable (category I, 3/18) or reduced (category II, 7/18) expression to that of controls. In contrast, 4/18 dogs displayed aberrant large vesicles revealed by the C-terminal but not the N-terminal antibody (category III), while 4/18 showed a control-like N-terminal expression but lacked the C-terminal protein (category IV). The missing C-terminal filaggrin in category IV strongly points towards loss-of function mutations in 4/18 (22%) of all AD dogs analysed.

**Key words:** atopic dermatitis – dog – filaggrin – hyperproliferation – nuclear factor-kappaB

*Accepted for publication 31 March 2010. Please cite this paper as*: Missing C-terminal filaggrin expression, NFkappaB activation and hyperproliferation identify the dog as a putative model to study epidermal dysfunction in atopic dermatitis. Experimental Dermatology 2010; **19**: e343–e346.

## Background

Filaggrin is a key molecule in epidermal differentiation, the formation of the cornified envelope (CE) and establishment of epidermal barrier function (1). In line with increased transepidermal water loss (TEWL), disturbed differentiation and epidermal hyperproliferation (2,3), loss-of-function mutations in the filaggrin gene (*FLG*) have recently been associated with human AD (4–8). Furthermore, activation of nuclear factor-kappaB (NFKB), which is one of the first responses of keratinocytes to alterations in epidermal barrier function and is often associated with increased epidermal proliferation, is considered a critical event in the progression and maintenance of AD (9,10).

Clinical features, histological lesions and immunologic aspects of canine AD show remarkable similarity to human AD (11,12). Reduced expression of terminal differentiation markers and increased TEWL consistent with impaired epidermal barrier function are also features of canine AD. This let us to hypothesize that, similar to human AD, canine AD might be associated with mutations in the *flg* gene, which is highly homologous to its human orthologue (http://www.ensemble.com). To the best of our knowledge, *flg* mutations have not been reported so far in canine AD (Appendix S1, suppl. information), and the expression pattern of other CE proteins such as involucrin and loricrin as well as NFKB activation and hyperproliferation have also not been addressed collectively in this canine disorder.

**Abbreviations:** AD, atopic dermatitis; FLG, filaggrin; NFKB, nuclear factor-kappaB.



## Question addressed

The more than 20 different *FLG* loss-of-function mutations described in human patients with AD invariably introduce a premature termination codon, as a result of nonsense or frameshift mutations, resulting in a C-terminal truncation of the filaggrin protein (1,8,13). Establishing a diagnostic sequencing strategy for human *FLG* has been a difficult task because of ten almost identical filaggrin repeats and two half-repeats encoding the >400 kDa profilaggrin protein (13). Hence, the possibility of *flg* mutations resulting in C-terminal protein truncations in canine patients with AD was addressed by other means. We raised a monospecific polyclonal C-terminal antibody to canine filaggrin and screened various breeds by immunofluorescence microscopy for lack of C-terminal versus presence of N-terminal filaggrin protein. We further linked these investigations to comprehensive analyses of CE, proliferation (Ki67 expression) and differentiation markers.

## Results

Immunofluorescence microscopy consistently revealed two distinct lines for both, the N-terminal and the C-terminal filaggrin antibody on all biopsies of healthy dogs ($n$ = 16) (Figs 1 and S1, suppl. methods). In comparison with haematoxylin and eosin–stained sections (Fig. S2), the inner (i.e. lower) line localized to the stratum granulosum and the outer (upper) to the bottom of the stratum corneum. In analogy to humans, this suggests that the inner line with a granular aspect (Figs 1 (C) and S1; insets) corresponds to immature pro-filaggrin in keratohyalin granules, while the outer line represents mature filaggrin. In addition to this pattern, the N-terminal antibody exhibited strong nuclear labelling consistent with nuclear translocation of the N-terminal filaggrin peptide described in humans (14). In non-lesional skin of 18 AD dogs of various breeds and mixed genders (see Appendix S1), the C-terminal antibody revealed four different highly distinctive filaggrin expression patterns (Fig. 1, I–IV C-terminal filaggrin). Category I was characterized by control-like filaggrin expression levels (3/18), which were visibly lower in categories II (7/18) and III (4/18). The distinctive peculiarity of category III, also observable by haematoxylin and eosin staining (Fig. S2), was the presence of large aberrant cytoplasmic vesicles in non-lesional skin (Fig. 1, III, insert). In lesional skin with a stronger hyperplasia, the number of epidermal layers containing these granules and their size further increased (Fig. 1, $III_L$). Like the C-terminal antibody, the N-terminal antibody revealed reduced filaggrin expression levels for categories II and III (Fig. 1). However, the large vesicles seen with the C-terminal antibody in category III were not visible with the N-terminal antibody (Fig. 1, III and $III^L$, inserts). Most strikingly, dogs of category IV (4/18) lacked specific immunolabelling for C-terminal filaggrin in both, non-lesional and lesional skin, but exhibited a strong control-like filaggrin expression with the N-terminal antibody (Fig. 1, IV shows non-lesional skin). Together, these data demonstrate that dogs of category I and II display the same expression pattern for the N- and C-terminal antibody, with overall weaker expression levels in category II. In contrast, dogs of categories III and IV exhibited differences in the staining pattern between the two antibodies. The accumulation of C-terminal but not N-terminal filaggrin in large vesicles may indicate abnormalities in the maturation and/or transport of the C-terminal filaggrin peptide. However, the detection of N-terminal but not C-terminal protein in category IV (4/18) highly supports the possibility of a loss-of-function mutation resulting in a C-terminal filaggrin truncation, analogous to those seen in humans.

Besides filaggrin, the expression of other proteins of the CE such as loricrin and desmoglein-1 (an essential adhesion molecule for epidermal integrity (15)) but not involucrin, was found to be focally reduced in non-lesional skin of all categories (Figs 1, 2a and S3). Moreover, in the case of moderate to strong hyperplasia, in particular in lesional skin, loricrin and involucrin were prematurely expressed. However, while involucrin levels were aberrantly increased, loricrin was reduced (Fig. S3, $I_L$ and $III_L$).

Compatible with findings in human AD, NFKB and the proliferation marker Ki67 were significantly increased in non-lesional skin of all categories described here (Fig. 2). While these alterations can also be caused by an inflammatory imbalance, they may be a direct consequence of an affected barrier function that can be expected in case of missing or aberrant C-terminal filaggrin expression in categories III and IV.

## Conclusion

Our results indicate that abnormalities in filaggrin protein expression are common in AD dogs (15/18), of which 4/18 (22%) present with a lack of C-terminal filaggrin protein, suggesting loss-of-function mutations. The exclusion approach we have designed proves to be a valuable strategy to identify candidate subjects to be sequenced for *flg* loss-of-function mutations and subsequently screen for these mutations within large cohorts of AD dogs. Our study further extends the similarity between human and canine AD to altered expression of epidermal differentiation markers, hyperproliferation and activation of NFKB. These results strengthen that the dog could represent a valuable model to gain further insights into the molecular mechanism of human AD (11).



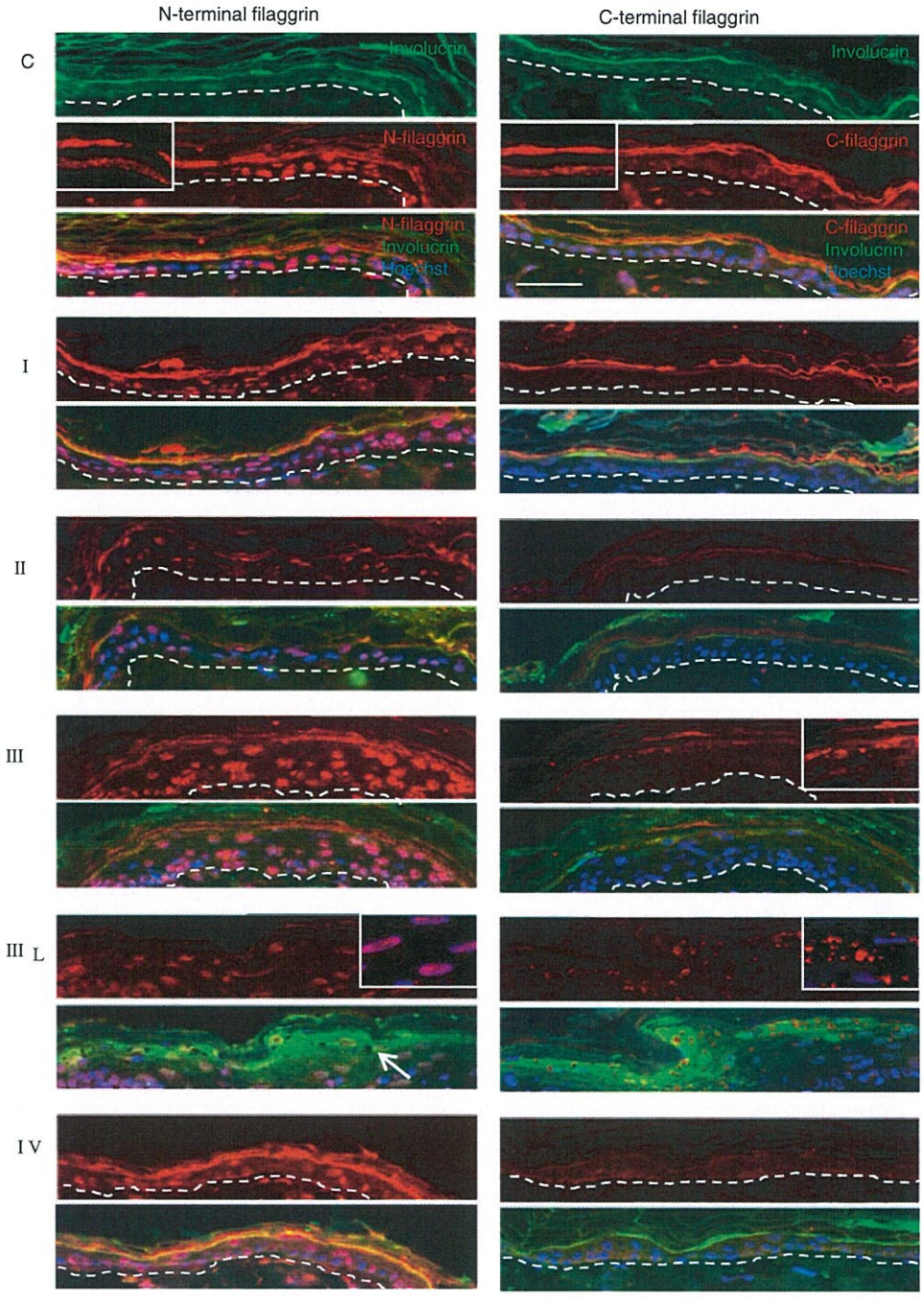



**Figure 1.** Filaggrin expression pattern in control and AD dogs. Shown are composite micrographs obtained by immunofluorescence microscopy of paraffin sections from control skin (C), non-lesional skin of AD dogs (categories I-IV, shows representative dogs) and lesional skin of one AD dog (category $III_L$) using the N-terminal (left panel) or C-terminal (right panel) filaggrin antibody. In all sections, immunolabelling for involucrin was merged with that for filaggrin (lower panel of each category). Inserts represent a three times higher magnification with adapted exposure time of a representative area to show the continuous and granular filaggrin lines, respectively, in C, cytoplasmic vesicles in III and their localization outside of nuclei (right) in contrast to the nuclear localization revealed by the N-terminal antibody (left) in $III_L$. Arrow points to gap in involucrin labelling at the location of vesicles stained with the C-terminal but not N-terminal antibody. Nuclei were counterstained with Hoechst 33258. Note that biopsies were processed simultaneously and photographic procedures held constant to allow comparison of intensities. Scale bar = 100 $\mu$m.

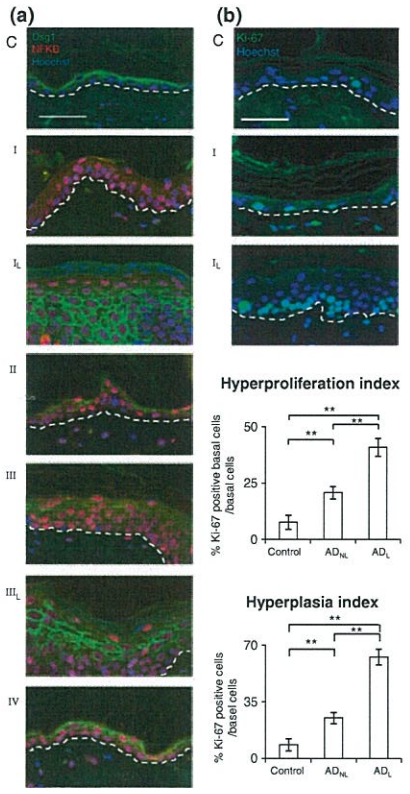



**Figure 2.** AD dogs exhibit increased NFKB activation, patchy desmoglein-1 expression and hyperproliferation as assessed by immunofluorescence microscopy on paraffin sections. (a) Immunolabelling for desmoglein-1 (green) and phospho-NFKB p65 (red) was merged. Note the general increase in NFKB and focally reduced desmoglein-1 expression in non-lesional skin of AD dogs. (b) Immunolabelling for Ki67 (green) on control, non-lesional and lesional skin. Shown are representative pictures. Biopsies of all dogs were stained and showed increased Ki67 expression (data not shown). For quantification, shown graphically, a subgroup of dogs was randomly selected, control; $n = 5$; AD non-lesional, $n = 6$ (representing all categories described in Fig. 2); AD lesional, $n = 3$. Categories and dogs are as in Fig. 1. Nuclei were counterstained with Hoechst 33258 (blue). The hyperproliferation (Ki67-positive basal cells over total number of basal cells) as well as the hyperplasia index (Ki67 positive over total number of basal cells) is depicted. Bars indicate mean ± SE. Data with significant $P$-values $^{**}P \leq 0.01$. Scale bars = 100 $\mu$m.

## Acknowledgements

Ludovic Chervet was supported by a grant from the Department of Clinical Veterinary Research. Filaggrin research in the McLean group is funded by grants from the Medical Research Council (reference G0700314), The British Skin Foundation and anonymous families affected by eczema in the Tayside region of Scotland. Huijia Chen is supported by a Scholarship from A*STAR, Singapore.

## Conflict of interest

Irwin McLean has filed patents relating to human filaggrin diagnosis and genetic testing.

## References

1 Sandilands A, Sutherland C, Irvine A D, McLean W H. J Cell Sci 2009: **122**: 1285–1294.
2 Proksch E, Folster-Holst R, Jensen J M. J Dermatol Sci 2006: **43**: 159–169.
3 Bieber T. N Engl J Med 2008: **358**: 1483–1494.
4 Palmer C N, Irvine A D, Terron-Kwiatkowski A *et al.* Nat Genet 2006: **38**: 441–446.
5 Smith F J, Irvine A D, Terron-Kwiatkowski A *et al.* Nat Genet 2006: **38**: 337–342.
6 Weidinger S, O'Sullivan M, Illig T *et al.* J Allergy Clin Immunol 2008: **121**: 1203–1209.
7 Weidinger S I T, Baurecht H, Irvine A D, Rodriguez E, Diaz-Lacava A, *et al.* J Allergy Clin Immunol 2006: **118**: 214–219.
8 Sandilands A, Terron-Kwiatkowski A, Hull P R *et al.* Nat Genet 2007: **39**: 650–654.
9 Dajee M, Muchamuel T, Schryver B *et al.* J Invest Dermatol 2006: **126**: 1792–1803.
10 Tanaka A, Muto S, Jung K, Itai A, Matsuda H. J Invest Dermatol 2007: **127**: 855–863.
11 Marsella R, Girolomoni G. J Invest Dermatol 2009: **129**: 2351–2357.
12 Marsella R, Olivry T. Clin Dermatol 2003: **21**: 122–133.
13 Brown S J, McLean W H. J Invest Dermatol 2009: **129**: 543–552.
14 O'Regan G M, Sandilands A, McLean W H, Irvine A D. J Allergy Clin Immunol 2009: **124**: R2–R6.
15 Amagai M, Matsuyoshi N, Wang Z H, Andl C, Stanley J R. Nat Med 2000: **6**: 1275–1277.

## Supporting Information

Additional Supporting Information may be found in the online version of this article:

**Figure S1.** Specificity of filaggrin antibodies. Filaggrin immunofluorescence staining on paraffin sections of a control dog using the pre-immune serum (from the rabbit injected with the C-terminal peptide), the C-terminal serum, and a commercially available N-terminal filaggrin antibody as indicated. Nuclei were counter-stained with Hoechst 33258. Inserts are three-fold magnifications of representative areas and illustrate a double filaggrin line. Note that the N-terminal antibody stains the nuclei.

**Figure S2.** Hematoxylin and Eosin (H&E) staining of representative sections from skin of humans and dogs from category I to IV. Controls (C) show overviews of normal human and canine skin. Of note is that healthy canine epidermis is thinner than human epidermis but has the same overall architecture. In both, normal human and canine skin, the keratohyalin granules are located in the stratum granulosum underneath the stratum corneum in a discontinuous line. Inserts are three times enlargements of a selected area showing keratohyalin granules. (I–IV) H&E sections corresponding to the four filaggrin categories described by means of immunofluorescence microscopy (Fig. 1). Results obtained in immunofluorescence microscopy reflected the situation observed on H&E sections; dogs of categories I and IV showed a similar distribution pattern and amount of keratohyalin granules to controls, whereas canine patients of category II had less detectable granules. In contrast, dogs of category III exhibited larger granules which were even bigger and expressed over several cell layers in lesional skin. Right panel is a 2.5 fold magnification of left panel. Scale bars = 100 $\mu$m.

**Figure S3.** Co-staining for loricrin and involucrin assessed by immunofluorescence microscopy on paraffin sections. Staining for involucrin and loricrin are shown separately or were merged as indicated for control (C), non-lesional (I–IV) and lesional skin (IL, IIIL). Exposure time was kept constant except for micrographs with insert. Inserts represent the same exposure time as the other micrographs; while micrographs IL, IIIL in the left panel were reduced and IL in the middle panel increased to visualize the staining pattern. Nuclei were counter-stained with Hoechst 33258. The same dogs as depicted in Fig. 1 are shown. Scale bar = 100 $\mu$m.

**Appendix S1.** Supplementary information, Material and Methods.

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.

*Original article*

# Raman profiles of the stratum corneum define 3 filaggrin genotype–determined atopic dermatitis endophenotypes

Gráinne M. O'Regan, MRCPI,[a,b] Patrick M. J. H. Kemperman, MD,[c] Aileen Sandilands, PhD,[d] Huijia Chen, MD,[d] Linda E. Campbell, BSc,[d] Karin Kroboth, PhD,[d] Rosemarie Watson, MD,[b] Marion Rowland, MB, PhD,[a,e] Gerwin J. Puppels, PhD,[f,g] W. H. Irwin McLean, DSc, FMedSci,[d] Peter J. Caspers, PhD,[f,g]* and Alan D. Irvine, MD, FRCPI[a,b,h]* *Dublin, Ireland, Rotterdam, The Netherlands, and Dundee, United Kingdom*

**Background:** Filaggrin (*FLG*) has a central role in the pathogenesis of atopic dermatitis (AD). *FLG* is a complex repetitive gene; highly population-specific mutations and multiple rare mutations make routine genotyping complex. Furthermore, the mechanistic pathways through which mutations in *FLG* predispose to AD are unclear.

**Objectives:** We sought to determine whether specific Raman microspectroscopic natural moisturizing factor (NMF) signatures of the stratum corneum could be used as markers of *FLG* genotype in patients with moderate-to-severe AD.

**Methods:** The composition and function of the stratum corneum in 132 well-characterized patients with moderate-to-severe AD were assessed by means of confocal Raman microspectroscopy and measurement of transepidermal water loss (TEWL). These parameters were compared with *FLG* genotype and clinical assessment.

**Results:** Three subpopulations closely corresponding with *FLG* genotype were identified by using Raman spectroscopy. The Raman signature of NMF discriminated between *FLG*-associated AD and non–*FLG*-associated AD (area under the curve, 0.94; 95% CI, 0.91-0.99). In addition, within the subset of *FLG*-associated AD, NMF distinguished between patients with 1 versus 2 mutations. Five novel *FLG* mutations were found on rescreening outlying patients with Raman signatures suggestive of undetected mutations (R3418X, G1138X, S1040X, 10085delC, and L2933X). TEWL did not associate with *FLG* genotype subgroups.

**Conclusions:** Raman spectroscopy permits rapid and highly accurate stratification of *FLG*-associated AD. *FLG* mutations do not influence TEWL within established moderate-to-severe AD. (J Allergy Clin Immunol 2010;■■■:■■■-■■■.)

*Key words: Atopic dermatitis, confocal Raman spectroscopy, eczema, filaggrin, hyperlinearity, natural moisturizing factor, transepidermal water loss, tyrosine*

Atopic dermatitis (AD) is a complex and heterogeneous inflammatory skin disease driven and modified by immunologic, environmental, and genetic factors.[1,2] The identification of filaggrin (*FLG*) null alleles in up to 50% of patients with moderate-to-severe AD implicates a fundamental role for barrier homeostasis in this disease.[3-6] Although the mechanisms leading to AD in *FLG* mutation carriers are unclear, the deficiency of *FLG* likely facilitates permeability of biologically active allergens and microbial colonization that subsequently trigger inflammatory cascades.[7] The recent identification of a murine model for *FLG* deficiency, with the detection of a homozygous frameshift mutation in the *Flg* gene in *flaky tail* mice, should accelerate our understanding of pathogenic mechanisms and therapeutic intervention points in patients with AD.[8]

Knowledge of the biochemical functions of filaggrin and its breakdown products indicate that a quantitative variation in gene or protein dosage might be relevant in determining phenotype.[9,10] Filaggrin is initially produced as profilaggrin, a large, insoluble, heavily phosphorylated protein consisting of 10 to 12 tandem repeats of filaggrin units separated by short hydrophobic linker peptides.[11,12] In the transitional layer profilaggrin is dephosphorylated and proteolytically processed into its functional filaggrin units, which bind to and collapse the keratin cytoskeleton and other intermediate filaments, acting as a scaffold for the subsequent reinforcement steps of the stratum corneum (SC).[13-15] Subsequently, filaggrin is progressively degraded within the SC into a pool of hygroscopic amino acids, including pyrrolidone carboxylic acid, urocanic acid, and alanine. This composite mixture of amino acids and their derivatives, together with specific salts and sugars, form the natural moisturizing factor (NMF).[16,17]

NMF is highly hygroscopic and plays a central role in maintaining hydration of the SC and is additionally proposed to have a significant role in maintenance of the pH gradient of the skin, cutaneous antimicrobial defense, and regulation of key enzymatic events in the SC.[18,19] Filaggrin thus has complex functions, with roles in establishing structural and chemical barrier function, hydration, and maintenance of epidermal homeostasis

From [a]the National Children's Research Centre and [b]the Department of Paediatric Dermatology, Our Lady's Children's Hospital, Dublin; [c]the Department of Dermatology and Venereology, Erasmus MC, Rotterdam; [d]the Epithelial Genetics Group, Medical Sciences Institute, University of Dundee; [e]the UCD School of Medicine and Medical Sciences, Dublin; [f]the Center for Optical Diagnostics and Therapy, Department of Dermatology and Venereology, Erasmus MC, Rotterdam; [g]River Diagnostics BV, Rotterdam; and [h]the Department of Medicine, Trinity College, Dublin.

*These authors contributed equally to this work.

Filaggrin research in the McLean laboratory is funded by grants from the British Skin Foundation, Medical Research Council, and donations from families affected by atopic dermatitis in the Tayside Region of Scotland. G.M.O'R. and A.D.I. are supported by the Children's Medical and Research Foundation, Dublin, Ireland.

Disclosure of potential conflict of interest: W. H. I. McLean receives research support from the British Skin Foundation. A. D. Irvine receives research support from the Children's Medical and Research Foundation. The rest of the authors have declared that they have no conflict of interest.

Received for publication December 7, 2009; revised April 27, 2010; accepted for publication April 29, 2010.

Reprint requests: Alan D. Irvine, MD, FRCPI, Department of Paediatric Dermatology, Our Lady's Children's Hospital, Crumlin, Dublin 12, Ireland. E-mail: irvinea@tcd.ie.

0091-6749/$36.00



doi:10.1016/j.jaci.2010.04.038

*Abbreviations used*
AD: Atopic dermatitis
AUC: Area under the curve
*FLG*: Filaggrin
HLP: Hyperlinear palms
NMF: Natural moisturizing factor
ROC: Receiver operating characteristic
SC: Stratum corneum
TEWL: Transepidermal water loss

in the face of continuous transformation.[20] Expression of filaggrin and subsequent breakdown of filaggrin into NMF is additionally determined based on properties of the microenvironment, including local pH, relative humidity, and protease activity.[17,21,22] *In vitro* evidence also indicates that filaggrin skin expression might be modulated by the atopic inflammatory response mediated by the cytokines IL-4 and IL-13.[23] Genetically determined modifiers of protein dosage include the copy number of filaggrin repeat units, which vary in the population from 10 to 12 units and segregate by normal Mendelian genetic mechanisms.[11,24] These genetic polymorphisms reflect tandem duplications of *FLG* repeats 8, 10, or both and might be an additional modifier of disease phenotype in heterozygotes who carry longer-sized variants on the unaffected allele.[25] Our early data suggest that NMF levels correlate with *FLG*-null allele status and might therefore directly contribute to the dry skin phenotype seen in both patients with ichthyosis vulgaris and those with AD.[10]

Raman spectroscopy is capable of measuring *in vivo* information regarding the molecular composition of the skin, including quantitative analysis of amino acids and water content. It is based on the inelastic light scattering, or Raman scattering, of monochromatic light when the frequency of photons, usually from a laser source, changes on interaction with a sample, giving rise to characteristic Raman spectra and providing noninvasive real-time signatures of biological samples at a molecular level.

We sought to determine whether specific Raman NMF signatures of the SC could be used as markers of *FLG* genotype in patients with moderate-to-severe AD. We examined the association of NMF estimation with clinical evidence of hyperlinear palms (HLP), a clinical sign that has been shown to be associated with *FLG* mutations in previous studies. We also sought to examine the effects, within patients with moderate-to-severe AD, of *FLG* genotype on transepidermal water loss (TEWL) as a measure of an inside-out barrier defect.

## METHODS

Following standard genetic practice, in this article $FLG^{-/-}$ designates a patient homozygous for null alleles (ie, 2 null alleles), $FLG^{+/-}$ designates a heterozygote null allele/wild-type (ie, 1 null allele), and $FLG^{+/+}$ designates a homozygote wild-type (ie, 0 null alleles). A further abbreviation describes patients with AD with *FLG* mutations ($FLG^{+/-}$ and $FLG^{-/-}$) as $AD_{FLG}$ and those without *FLG* mutations (ie, $FLG^{+/+}$) as $AD_{NON\text{-}FLG}$.

One hundred thirty-five unrelated Irish children with a history of moderate-to-severe AD were recruited from dedicated tertiary referral AD clinics. Diagnosis was made by experienced pediatric dermatologists according to the United Kingdom diagnostic criteria.[26] Exclusion criteria from the study were patients who had received systemic therapy, such as corticosteroids or immunosuppressants, in the preceding 3 months and patients whose ancestry was not exclusively Irish (4/4 grandparents). Detailed phenotypic data were collected. The Nottingham Eczema Severity Score[27] was selected as an estimate of disease severity. Given previous publications and our own clinical experience of noting palmar hyperlinearity, this clinical sign was scored by using an investigator assessment of 0 (no hyperlinearity), 1 (mild/subtle hyperlinearity), and 2 (severe hyperlinearity).

### Genetic screening

All patients were screened for the 6 most prevalent *FLG* mutations in the Irish population (R501X, 2282del4, R2447X, S3247X, 3702delG, and Y2092X), as previously described.[28] Full sequencing of *FLG* was performed as detailed previously.[28] Based on screening for these 6 prevalent *FLG* mutations, 58.3% were carriers of 1 or more *FLG* mutations (15.1% $FLG^{-/-}$, 43.2% $FLG^{+/-}$, and 41.7% $FLG^{+/+}$). An additional rare mutation was known to be present in 1 subject in a heterozygous state (R1474X). Additional screening was performed by means of complete sequencing of the *FLG* gene in selected subjects, as previously described.[29] The entire collection was then rescreened for these 5 novel mutations found on complete sequencing.

### Biophysical analysis of the SC

Skin biophysical measurements were performed under standardized conditions (room temperature, 22 °C-25 °C; humidity levels, 30% to 35%). Before measurements, patients were acclimatized for a minimum of 10 minutes. All measurements were performed by one of 2 investigators (G.M.O. and P.M.J.H.K.). Topical therapies, including emollients, were withheld from the measurement sites for 48 hours preceding the study. TEWL was measured on nonlesional skin of the extensor forearm (Tewameter 300; Courage and Khazaka Electronic GmbH, Cologne, Germany).

NMF was measured in the SC of the thenar eminence by using confocal Raman microspectroscopy (model 3510 Skin Composition Analyzer; River Diagnostics, Rotterdam, The Netherlands). The principles of this method and the procedure have been described elsewhere.[30,31] Depth profiles of Raman spectra were measured at 5-μm intervals from the skin surface at 30, 35, 40, 45 and 50 μm below the skin surface. An average of 8 profiles (totaling to 40 Raman spectra) from different areas of the thenar eminence were measured per patient. Raman spectra were recorded in the spectral region at 400 to 1,800 $cm^{-1}$ with a 785-nm laser. Laser power on the skin was 25 mW. Levels of skin constituents relative to keratin were determined from the Raman spectra by means of classical least-squares fitting. Details of the method have been described elsewhere.[30,31] Briefly, reference spectra of keratin, NMF, urocanic acid, lactate, urea, ceramide, and cholesterol were fitted to the individual Raman spectra from the skin. The combination of these reference spectra provides an adequate model for *in vivo* Raman spectra of normal human SC. A spectrum of reagent-grade L-tyrosine (Sigma-Aldrich, Zwijndrecht, The Netherlands) was added to this set of reference-fit spectra to enable determination of increased tyrosine levels. The resulting fit coefficients represent the relative proportions in which the skin constituents contribute to the total Raman skin spectrum. The reference spectrum of NMF had been constructed from the weighted sum of the spectra of its dominant constituents (pyrrolidone carboxylic acid, ornithine, serine, proline, glycine, histidine, and alanine). All concentrations of skin constituents relative to keratin were calculated from the recorded Raman spectra by using SkinTools 2.0 (River Diagnostics B.V., Rotterdam, The Netherlands); NMF levels derived from the individual Raman spectra were used to assess intrapatient variation in NMF. These NMF levels were then averaged to obtain the mean NMF level per patient. Of the 135 subjects who participated in the study, 3 patients were unable to fully cooperate with the Raman measurement, and their results were excluded because of insufficient spectra collection.

The study was conducted in accordance with the Declaration of Helsinki and was approved by the Research Ethics Committee of Our Lady's Children's Hospital, Dublin. Written informed consent was obtained from all patients or their parents.

### Statistical methods

Patients were characterized, *a priori*, into 3 genotypes ($FLG^{+/+}$, $FLG^{+/-}$, and $FLG^{-/-}$), as described in the methods. The designation *FLG-associated AD* ($AD_{FLG}$) includes those with 1 or 2 *FLG* mutations ($FLG^{+/-}$ and $FLG^{-/-}$),

whereas non–*FLG*-associated AD ($AD_{NON\text{-}FLG}$) are $FLG^{+/+}$. The data were characterized as having either a normal or a skewed distribution. Positively skewed data were log transformed. Box and whiskers plots showing the median value with interquartile range (25th-75th box length) of variables were constructed for the 3 *FLG* genotypes ($FLG^{-/-}$, $FLG^{+/-}$, and $FLG^{+/+}$). ANOVA was used to compare means among the 3 genotype subgroups and presented as means with SDs and 95% CIs. Homogeneity of variances was tested by using the Levine statistic. A *post hoc* Tukey analysis for multiple comparisons was performed to examine pairwise differences among the 3 genotype subgroups. Mean differences for each pairwise comparison was presented together with a 95% CI of the mean difference.

Nonparametric receiver operating characteristic (ROC) curves, which plot sensitivity against 1−specificity, were constructed to examine the utility of Raman spectroscopy to, in the first instance, differentiate between $AD_{FLG}$ and $AD_{NON\text{-}FLG}$. We then sought to further determine whether, within the *FLG*-associated AD group, ROC curve analysis could determine an NMF cutoff point to predict homozygous or heterozygous subjects ($FLG^{-/-}$ vs $FLG^{+/-}$). Areas under the curve (AUCs) with 95% CIs are presented. The cutoff for maximal sensitivity and specificity was deduced from the graphic output. Sensitivity and specificity for our dataset, based on the ROC output, are also presented to provide a preliminary indication of the clinical utility of Raman spectroscopy as a novel technique to distinguish *FLG* genotypes in patients with AD. However, further evaluation of Raman spectroscopy will be required in different well-characterized populations of patients with AD and healthy subjects.

HLP was scored clinically as absent (0) mild (1), or severe (2). For statistical analysis, HLP was dichotomized to no evidence of hyperlinearity (clinical score 0) versus any hyperlinearity (clinical scores of 1 or 2). The κ statistic was used to measure agreement between clinical scores of HLP and an NMF cutoff of 1.07. Where the observed agreement between scores is better than the degree of agreement expected by chance alone, κ is scored from 0 to 1.0, with 1.0 representing perfect agreement. A κ value of less than 0.2 is poor, a score of 0.2 to 0.4 is fair, 0.4 to 0.6 is considered moderate agreement, 0.6 to 0.8 is considered good agreement, and greater than 0.8 is considered excellent agreement.[32] Data were analyzed with SPSS software (version 15; SPSS, Inc, Chicago, Ill). Significance was set at the 5% level.

## RESULTS

Clinical characteristics and summary data of the study cohort are outlined in Table I. NMF levels assessed by means of Raman spectroscopy for subjects according to final *FLG* genotype (final genotype after full screening) are shown in Fig 1. In each pairwise comparison of the 3 genotypes, there is a statistically significant difference in NMF values ($P < .001$, ANOVA, Tukey post hoc analysis; Table II). IgE levels were not normally distributed. After log transformation of IgE data, there was no statistically significant difference in IgE values among the 3 *FLG* mutation subgroups ($P < .48$, Table I).

Within the $FLG^{+/+}$ and $FLG^{+/-}$ genotype subgroups, 13 subjects had significantly outlying values with significantly lower NMF values compared with others who shared these genotypes (Fig 2). We hypothesized that these outlying values were suggestive of additional undetected *FLG* mutations. These 13 subjects were therefore fully sequenced for potential additional mutations. After full sequencing of these 13 subjects, 5 novel mutations (R3418X, G1138X, S1040X, 10085delC and L2933X) were detected. The entire collection was then rescreened to include 11 mutations (6 on initial screening plus these 5 additional mutations); however, no additional patients were found to harbor these novel rare mutations. NMF values by genotype after this additional screening are plotted in histogram format in Fig 3. Within *FLG* genotype subgroups, a range of values was seen; these values approximate to a normal distribution in each of the 3 genotype subgroups. The new mutations increased the percentage of patients with 1 or more *FLG* mutations in the collection to 59.8% (18.2% $FLG^{-/-}$, 41.7% $FLG^{+/-}$, and 40.2% $FLG^{+/+}$; Table I).

### NMF values predict *FLG* genotype in patients with moderate-to-severe AD

ROC curves were constructed to examine the discriminatory power of Raman-determined NMF. From a clinical decision tree, patients were initially grouped as $AD_{FLG}$ and $AD_{NON\text{-}FLG}$ ($FLG^{+/+}$). The discriminatory power of Raman-determined NMF here was high, with an AUC on ROC analysis of 0.95 (95% CI, 0.91-0.99; Fig 4). The optimal cutoff value for NMF was 1.07, which would equate to a sensitivity of 98.73% and a specificity of 86.89% extrapolated from our ROC analysis. With the $AD_{FLG}$ group of children, NMF values also distinguished between $FLG^{+/-}$ and $FLG^{-/-}$ subjects, with an ROC AUC of 0.85 (95% CI, 0.77-0.93; see Fig E1 in this article's Online Repository at www.jacionline.org). This sensitivity and specificity is presented to suggest the probable utility of Raman spectroscopy as a novel diagnostic technique in patients with AD and is not presented as a test for the goodness of fit of the model.

### Tyrosine peaks on Raman microspectroscopy predict $FLG^{-/-}$ genotype

The Raman spectra collected in the study revealed a distinct signal profile in a subset of patients subsequently identified by matching of *in vitro* Raman spectra of the amino acid tyrosine. Tyrosine levels 5 SDs greater than the normal tyrosine level measured in $FLG^{+/+}$ subjects were identified as increased. Above this threshold, distinct tyrosine peaks could be visually identified in the corresponding Raman spectra. Tyrosine peaks were present in 19 (79%) of 24 of the patients who had 2 *FLG* mutations (Table I) and absent from all but 1 $FLG^{+/+}$ subject, who also had an outlying low NMF level (Fig 5), suggesting an undetected genetic defect in *FLG*. The optimum NMF cutoff to correlate with tyrosine peaks was 0.831 (AUC, 0.81; 95% CI, 0.73-0.89), which is similar to the cutoff that distinguishes the $FLG^{-/-}$ from the $FLG^{+/-}$ populations (data not shown).

### TEWL in patients with moderate-to-severe AD does not discriminate between *FLG* genotype subgroups

There was no difference in mean TEWL levels among the 3 *FLG* genotype subgroups (Table I). TEWL did not distinguish *FLG* genotype status in patients with AD (TEWL AUC, 0.583; 95% CI, 0.48-0.699; ROC curve not shown).

### NMF values and palmar hyperlinearity scores

There was good agreement between clinical assessment of palmar hyperlinearity and an NMF cutoff of 1.07 ($\kappa = 0.71$). Thus palmar hyperlinearity might help distinguish patients with AD into those with 1 or more *FLG* mutations versus those with none. Clinical assessment is not as accurate as Raman spectroscopy in distinguishing the 3 *FLG* genotypes. The Nottingham Eczema Severity Score did not segregate with *FLG* genotype in this collection (data not shown), suggesting that the spectrum of AD severity within this moderate-to-severe collection is independent of *FLG* genotype.

**TABLE I.** Cohort characteristics according to final genotype

| Final *FLG* genotype*† | All mutations combined, no. (%) | Age (y), mean (SD) | Male sex, no. (%) | NESS,‡ mean (SD) | Palmar hyperlinearity score‡ | TEWL (g/m²/h), mean (SD) | NMF (AU), mean (SD) | Tyrosine peaks, no. (%) | Log IgE# mean (SD) |
|---|---|---|---|---|---|---|---|---|---|
| Screened§ | n = 132 | n = 132 | n = 132 | n = 132 | n = 131 | n = 102 | n = 132 | n = 132 | n = 129 |
| +/+ | 53 (40.15) | 8.43 (3.76) | 32 (60.3) | 12.40 (2.39) | 0: 36 | 15.54 (9.34), n = 41 | 1.31 (0.24) | 1 (1.89) | 6.89 (2.29) |
| | | | | | 1: 10 | | | | |
| | | | | | 2: 7 | | | | |
| | | | | | n = 53 | | | | |
| +/− | 55 (41.66) | 8.2 (4.12) | 32 (58.2) | 11.28 (2.80) | 0: 8 | 15.59 (6.72), n = 44 | 0.80 (0.21) | 10 (18.18) | 6.56 (1.78) |
| | | | | | 1: 23 | | | | |
| | | | | | 2: 23 | | | | |
| | | | | | n = 54 | | | | |
| −/− | 24 (18.18) | 8.83 (4.16) | 15 (62.5) | 12.13 (2.52) | 0: 0 | 17.35 (7.37), n = 17 | 0.57 (0.15) | 19 (79.16) | 7.11 (1.50) |
| | | | | | 1: 1 | | | | |
| | | | | | 2: 23 | | | | |
| | | | | | n=24 | | | | |
| *P* value | – | .84‖ | .97¶ | .07‖ | – | .70‖ | <.0001‖ | <.0001¶ | .48‖ |

*AU*, Arbitrary units.
*Final *FLG* status after rescreening cohort for all mutations (*FLG* status before detection of novel mutations: $FLG^{+/+}$, 41.7%; $FLG^{+/-}$, 43.2%; and $FLG^{-/-}$, 15.1%).
†+/+, No *FLG* mutation; +/−, 1 *FLG* mutation; −/−, 2 *FLG* mutations.
‡*NESS*, Nottingham Eczema Severity Score. Palmar hyperlinearity score: 0, no hyperlinearity; 1, intermediate; 2, marked.
§Hyperlinearity scoring and TEWL data are not available for 1 and 30 subjects, respectively.
‖Comparison of means among the 3 groups of *FLG* mutations using ANOVA.
¶Comparison of proportions using $\chi^2$ test for comparison of a 2 × 3 contingency table.
#IgE data were positively skewed. The mean of the log-transformed data is presented.

## DISCUSSION

A primary aim of this study was to investigate the utility of *in vivo* measurements of the SC to differentiate between *FLG* genotypes within patients with AD. Second, by making use of a very well-genotyped collection, we sought to determine whether *FLG* genotype influenced TEWL in patients with established moderate-to-severe AD. These results establish that the biochemical composition of the SC and palmar hyperlinearity associate with *FLG* mutation status, allowing rapid stratification of AD endotypes. Furthermore, we conclude that the degree of barrier defect, as measured by TEWL, in the clinically unaffected skin of patients with established AD is not influenced by *FLG* genotype.

The clinical utility of Raman-determined NMF is powerfully illustrated in the ROC curve analysis, which has an AUC of 0.95 (95% CI, 0.91-0.99) to discriminate between the $AD_{FLG}$ and $AD_{NON\text{-}FLG}$ groups. In addition, NMF can distinguish from heterozygous *FLG* mutations. In other words, this technique allows segregation of AD into *FLG* genotypes (+/+, +/−, and −/−) with such high specificity that it can accurately predict *FLG* genotype. Even in this relatively small patient population, mean NMF was statistically different between each of the 3 subpopulations (*post hoc* Tukey, Table II) Further validation of Raman-determined NMF is required in patients with well-characterized AD and healthy populations before this novel technique can be considered for clinical use.

The observed interindividual variation in NMF results within the mutation groups approximates to a normal distribution (Fig 3) that might be caused by environmental, genetic, or immunologic modifiers of *FLG* expression, including posttranslational processing. Homeostatic mechanisms or genetic mechanisms, such as the effects of copy number variation, small gene deletions, cytokine modulation of *FLG* expression, or upregulation of additional sources of NMF constituents, such as eccrine-derived lactate and glycerol, might account for differences in NMF levels within genotype groups. Modifying genes involved in filaggrin-processing pathways are likely to additionally contribute to this variation.[20] The strong relationship between *FLG* mutation status and NMF implies that the role of other histidine-rich barrier proteins, such as the S100-fused proteins, including filaggrin-2 and hornerin, that are thought to have synergistic roles with filaggrin, including contribution to NMF levels,[33-35] are unlikely to play a major role in the total NMF seen in AD patients.

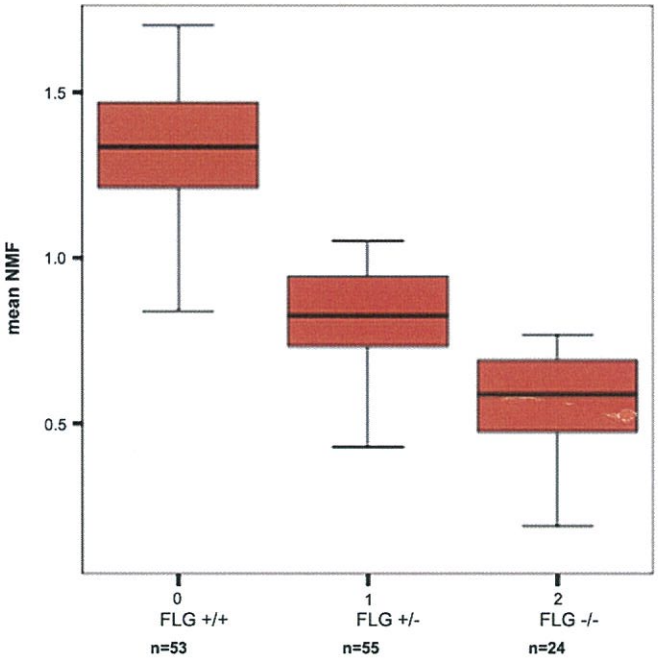



**FIG 1.** Box and whiskers plot of NMF by *FLG* genotypes (final genotype after full screening) showing the median *(midline)* and interquartile range corresponding to the length of the box.

The strong predictive value of a tyrosine signal in the SC of subjects with *FLG* mutations identifies a potential biomarker in

**TABLE II.** ANOVA showing a statistically significant difference in NMF among the 3 *FLG* genotype subgroups together with 95% CIs

| Genotype | No. | Mean NMF | 95% CI | Comparison genotype | Mean difference | 95% CI mean difference | *P* value* |
|---|---|---|---|---|---|---|---|
| $FLG^{+/+}$ | 53 | 1.31 ± 0.24 | 1.24 to 1.37 | $FLG^{+/-}$ | 0.51 | 0.40 to 0.60 | <.001 |
| | | | | $FLG^{-/-}$ | 0.74 | 0.62 to 0.87 | <.001 |
| $FLG^{+/-}$ | 55 | 0.80 ± 0.21 | 0.74 to 0.86 | $FLG^{+/+}$ | −0.51 | −0.60- to −0.41 | <.001 |
| | | | | $FLG^{-/-}$ | 0.24 | 0.11 to 0.36 | <.001 |
| $FLG^{-/-}$ | 24 | 0.57 ± 0.15 | 0.50 to 0.63 | $FLG^{+/+}$ | −0.74 | −8.86 to −0.61 | <.001 |
| | | | | $FLG^{+/-}$ | −0.24 | −0.36 to −0.11 | <.001 |

A *post hoc* analysis with Tukey correction shows a statistically significant difference in mean NMF between each pairwise comparison for the 3 *FLG* genotype subgroups.
*Tukey multiple comparisons.

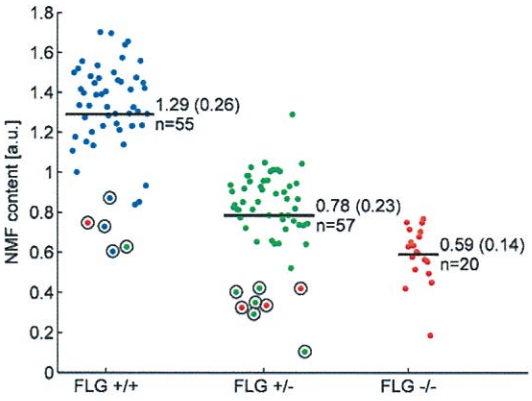



**FIG 2.** NMF cloud plot of values subcategorized after initial screening of 6 prevalent mutations. *Circles* indicate outliers who were rescreened for additional mutations. *Colors* indicate the final genotype after full screening. For each group, the number of patients and average NMF level (mean ± SD) are indicated in the figure. Group comparisons showing means and standard deviations by using ANOVA are presented in the text. *a.u.*, Arbitrary units.

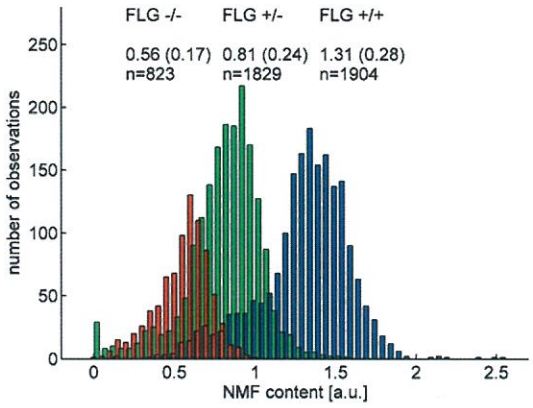



**FIG 3.** Histogram of all nonaveraged NMF values subcategorized according to corrected mutational status. *Blue bars* are $FLG^{+/+}$ subjects, *green bars* are $FLG^{+/-}$ subjects, and *red bars* are $FLG^{-/-}$ subjects. For each group, the number of Raman measurements and the average NMF level (mean ± SD) are indicated in the figure. *FLG* genotype group sizes are as follows: $FLG^{-/-}$, 24; $FLG^{+/-}$, 55; $FLG^{+/+}$, 53. *a.u.*, Arbitrary units.

the diagnosis of *FLG*-null subjects. Filaggrin lacks tyrosine; however, the linker segments and carboxyl-terminal domains of profilaggrin contain dense and highly conserved tyrosine-rich motifs.[11,12,36] In normal SC only tyrosine incorporated in proteins, predominantly keratin, contributes to the overall Raman spectrum of skin; unbound tyrosine is usually not detectable. The consistent observation of the presence of tyrosine peaks in homozygous *FLG* mutation carriers might relate to altered biosynthetic pathways, such as limited proteolysis of profilaggrin or aberrant intermediate processing species, degraded through alternative pathways. The presence of tyrosine is independent of disease severity or mutation location.

Palmar and plantar hyperlinearity is a consistent clinical finding in the monogenic *FLG* disease ichthyosis vulgaris, and in keeping with a semidominant trait, intermediate hyperlinearity is thought to be associated with heterozygosity for *FLG* mutations; hyperlinearity has been associated as a significant finding in patients with *FLG*-related AD.[37-40] Brown et al[41] have recently shown significant association of palmar hyperlinearity with *FLG*-null mutations in a population-based cohort.[41] Our finding that *FLG* status and NMF segregate with palmar hyperlinearity in patients with moderate-to-severe AD further validates the importance of this clinical sign in AD phenotypes.

Evidence from both molecular genetics and functional analyses strongly supports a skin barrier defect as a feature of both lesional and nonlesional skin in patients with AD.[42-44] Increased mean baseline TEWL levels, reflecting a dysfunction of the inside-out epidermal permeability barrier in the clinically uninvolved skin of patients with AD, have been shown in many previous studies.[45-47] The concept of an inherent skin barrier defect in both lesional and nonlesional skin from patients with is further supported by other methodologies, including recent data by Jakasa et al[43] showing enhanced uptake of entire series of polyethylene glycols covering molecular weights in the range 150 to 590 daltons in nonlesional skin of patients with AD. Hata et al, using a photoacoustic spectroscopic system, showed enhanced penetration of both lipophilic and hydrophilic dye through clinically normal skin of patients with AD compared with that seen in control subjects.[42-44]

Previous investigators have reported contrasting findings on the influence of *FLG* mutation status on the inside-out barrier, as measured by TEWL in patients with AD; however, these studies were restricted by small sample size and possibly confounded by the technical limitations of disclosure of full *FLG* polymorphism status in the populations studied.[10,48] A French study including a small number of *FLG* mutation carriers suggested that TEWL in patients with AD was not influenced by *FLG* status.[49] Here, in our much larger collection of comprehensively genotyped patients with moderate-to-severe AD, TEWL again showed no discernable difference between *FLG* genotype subpopulations. These results suggest that increased TEWL in patients with moderate-to-severe AD is independent of *FLG* status and is a common end point in patients with moderate-to-severe AD. Thus increased TEWL might result from the systemic effects of the inflammatory

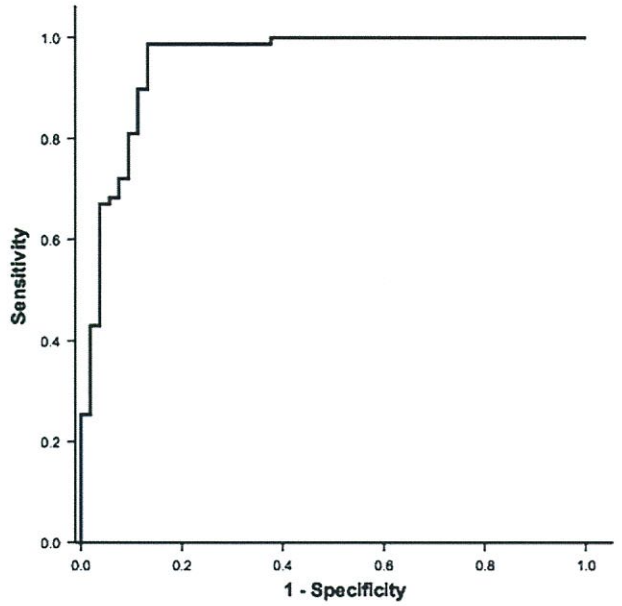



**FIG 4.** ROC curve for NMF $AD_{FLG}$ (genotype +/− and −/−) compared with $AD_{NON\text{-}FLG}$ (genotype +/+). The AUC is 0.95 (95% CI, 0.91-0.99). The optimal cutoff point for mean NMF to distinguish $AD_{FLG}$ from $AD_{NON\text{-}FLG}$ was 1.07 arbitrary units.

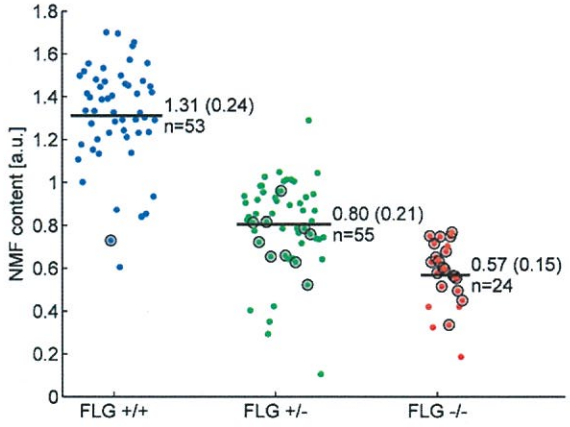



**FIG 5.** Cloud plot of NMF values categorized by genotype (final genotype after full screening). *Circles* indicate subjects with an increased tyrosine signal ($FLG^{+/+}$, n = 1; $FLG^{+/-}$, n = 10; $FLG^{-/-}$, n = 19). For each group, the number of patients and average NMF level (mean ± SD) are indicated in the figure. *a.u.*, Arbitrary units.

process, even at nonclinically inflamed sites. Interestingly, flaky tail *(ft)* mice show a moderate increase in baseline TEWL, a deviation that increases dramatically after allergen priming and the subsequent secondary inflammatory response.[8] In these mice, although *FLG* deficiency might of itself lead to an inside-out barrier defect barely measureable by increased TEWL, the subsequent marked increase in TEWL appears to be driven by the secondary immunologic response.

Within this collection, no effect was seen for *FLG* and disease severity or total serum IgE level; however, this case series is not the ideal collection to examine this because all cases are moderate to severe and therefore the full range of severity is not represented. Further population-based and sufficiently powered studies should clarify this potential genotype-phenotype correlation.

Within patients with AD, Raman signatures of NMF level and tyrosine predict *FLG* mutation status, overcoming the need for technically demanding genotyping, particularly in populations in which the *FLG* mutation architecture has not been comprehensively elucidated. This work additionally further informs mechanistic pathways in AD, demonstrating that TEWL in patients with moderate-to-severe AD is an outcome that is independent of *FLG* mutation status and pointing to a complex systemic interplay of environmental, immunologic, and functional pathways in the genesis of epithelial barrier disruption in patients with AD.

We thank the patients and their families for their involvement in this study.

**Key messages**

- **Raman spectroscopy, specifically NMF analysis, permits rapid and accurate stratification of FLG-associated versus non–FLG-associated AD and might be useful as a predictive test for FLG mutations.**
- **There is good agreement between FLG mutation status, NMF, and clinical palmar hyperlinearity scoring in patients with moderate-to-severe AD. Palmar hyperlinearity might help to define a clinical and genetic phenotype of AD.**
- **Within patients with moderate-to-severe AD, TEWL levels do not segregate with FLG genotype, suggesting that FLG mutations are neither necessary nor sufficient to fully explain this inside-out barrier defect in established AD.**

**REFERENCES**

1. Bieber T. Atopic dermatitis. N Engl J Med 2008;358:1483-94.
2. Leung DY, Bieber T. Atopic dermatitis. Lancet 2003;361:151-60.
3. Sandilands A, O'Regan GM, Liao H, Zhao Y, Terron-Kwiatkowski A, Watson RM, et al. Prevalent and rare mutations in the gene encoding filaggrin cause ichthyosis vulgaris and predispose individuals to atopic dermatitis. J Invest Dermatol 2006; 126:1770-5.
4. Barker JN, Palmer CN, Zhao Y, Liao H, Hull PR, Lee SP, et al. Null mutations in the filaggrin gene (FLG) determine major susceptibility to early-onset atopic dermatitis that persists into adulthood. J Invest Dermatol 2007;127:564-7.
5. Baurecht H, Irvine AD, Novak N, Illig T, Buhler B, Ring J, et al. Toward a major risk factor for atopic eczema: meta-analysis of filaggrin polymorphism data. J Allergy Clin Immunol 2007;120:1406-12.
6. Weidinger S, O'Sullivan M, Illig T, Baurecht H, Depner M, Rodriguez E, et al. Filaggrin mutations, atopic eczema, hay fever, and asthma in children. J Allergy Clin Immunol 2008;121:1203-9, e1.
7. Elias PM. Skin barrier function. Curr Allergy Asthma Rep 2008;8:299-305.
8. Fallon PG, Sasaki T, Sandilands A, Campbell LE, Saunders SP, Mangan NE, et al. A homozygous frameshift mutation in the mouse Flg gene facilitates enhanced percutaneous allergen priming. Nat Genet 2009;41:602-8.
9. Ginger RS, Blachford S, Rowland J, Rowson M, Harding CR. Filaggrin repeat number polymorphism is associated with a dry skin phenotype. Arch Dermatol Res 2005;297:235-41.
10. Kezic S, Kemperman PM, Koster ES, de Jongh CM, Thio HB, Campbell LE, et al. Loss-of-function mutations in the filaggrin gene lead to reduced level of natural moisturizing factor in the stratum corneum. J Invest Dermatol 2008;128:2117-9.
11. Gan SQ, McBride OW, Idler WW, Markova N, Steinert PM. Organization, structure, and polymorphisms of the human profilaggrin gene. Biochemistry 1990;29: 9432-40.
12. Presland RB, Haydock PV, Fleckman P, Nirunsuksiri W, Dale BA. Characterization of the human epidermal profilaggrin gene. Genomic organization and identification of an S-100-like calcium binding domain at the amino terminus. J Biol Chem 1992; 267:23772-81.
13. Segre J. Complex redundancy to build a simple epidermal permeability barrier. Curr Opin Cell Biol 2003;15:776-82.

14. Kalinin AE, Kajava AV, Steinert PM. Epithelial barrier function: assembly and structural features of the cornified cell envelope. Bioessays 2002;24:789-800.
15. Candi E, Schmidt R, Melino G. The cornified envelope: a model of cell death in the skin. Nat Rev Mol Cell Biol 2005;6:328-40.
16. Rawlings AV, Scott IR, Harding CR, Bowser PA. Stratum corneum moisturization at the molecular level. J Invest Dermatol 1994;103:731-41.
17. Scott IR, Harding CR. Filaggrin breakdown to water binding compounds during development of the rat stratum corneum is controlled by the water activity of the environment. Dev Biol 1986;115:84-92.
18. Rawlings AV, Matts PJ. Stratum corneum moisturization at the molecular level: an update in relation to the dry skin cycle. J Invest Dermatol 2005;124:1099-110.
19. Krien PM, Kermici M. Evidence for the existence of a self-regulated enzymatic process within the human stratum corneum—an unexpected role for urocanic acid. J Invest Dermatol 2000;115:414-20.
20. Sandilands A, Sutherland C, Irvine AD, McLean WH. Filaggrin in the frontline: role in skin barrier function and disease. J Cell Sci 2009;122:1285-94.
21. Bouwstra JA, Groenink HW, Kempenaar JA, Romeijn SG, Ponec M. Water distribution and natural moisturizer factor content in human skin equivalents are regulated by environmental relative humidity. J Invest Dermatol 2008;128:378-88.
22. Nakagawa N, Sakai S, Matsumoto M, Yamada K, Nagano M, Yuki T, et al. Relationship between NMF (lactate and potassium) content and the physical properties of the stratum corneum in healthy subjects. J Invest Dermatol 2004;122:755-63.
23. Howell MD, Kim BE, Gao P, Grant AV, Boguniewicz M, Debenedetto A, et al. Cytokine modulation of atopic dermatitis filaggrin skin expression. J Allergy Clin Immunol 2007;120:150-5.
24. McKinley-Grant LJ, Idler WW, Bernstein IA, Parry DA, Cannizzaro L, Croce CM, et al. Characterization of a cDNA clone encoding human filaggrin and localization of the gene to chromosome region 1q21. Proc Natl Acad Sci U S A 1989;86: 4848-52.
25. Wong KK, deLeeuw RJ, Dosanjh NS, Kimm LR, Cheng Z, Horsman DE, et al. A comprehensive analysis of common copy-number variations in the human genome. Am J Hum Genet 2007;80:91-104.
26. Williams HC, Burney PG, Hay RJ, Archer CB, Shipley MJ, Hunter JJ, et al. The U.K. Working Party's Diagnostic Criteria for Atopic Dermatitis. I. Derivation of a minimum set of discriminators for atopic dermatitis. Br J Dermatol 1994;131: 383-96.
27. Emerson RM, Charman CR, Williams HC. The Nottingham Eczema Severity Score: preliminary refinement of the Rajka and Langeland grading. Br J Dermatol 2000;142:288-97.
28. Sandilands A, Smith FJ, Irvine AD, McLean WH. Filaggrin's fuller figure: a glimpse into the genetic architecture of atopic dermatitis. J Invest Dermatol 2007;127:1282-4.
29. Sandilands A, Terron-Kwiatkowski A, Hull PR, O'Regan GM, Clayton TH, Watson RM, et al. Comprehensive analysis of the gene encoding filaggrin uncovers prevalent and rare mutations in ichthyosis vulgaris and atopic eczema. Nat Genet 2007;39:650-4.
30. Caspers PJ, Lucassen GW, Carter EA, Bruining HA, Puppels GJ. In vivo confocal Raman microspectroscopy of the skin: noninvasive determination of molecular concentration profiles. J Invest Dermatol 2001;116:434-42.
31. Caspers PJ, Lucassen GW, Puppels GJ. Combined in vivo confocal Raman spectroscopy and confocal microscopy of human skin. Biophys J 2003;85:572-80.
32. Daly LE, Bourke G. Interpretation and uses of medical statistics. 5th ed. Oxford: Blackwell Scientific Press; 2000. p. 381-421.
33. Esparza-Gordillo J, Weidinger S, Folster-Holst R, Bauerfeind A, Ruschendorf F, Patone G, et al. A common variant on chromosome 11q13 is associated with atopic dermatitis. Nat Genet 2009;41:596-601.
34. Wu Z, Hansmann B, Meyer-Hoffert U, Glaser R, Schroder JM. Molecular identification and expression analysis of filaggrin-2, a member of the S100 fused-type protein family. PLoS One 2009;4:e5227.
35. Toulza E, Mattiuzzo NR, Galliano MF, Jonca N, Dossat C, Jacob D, et al. Large-scale identification of human genes implicated in epidermal barrier function. Genome Biol 2007;8:R107.
36. Resing KA, Walsh KA, Haugen-Scofield J, Dale BA. Identification of proteolytic cleavage sites in the conversion of profilaggrin to filaggrin in mammalian epidermis. J Biol Chem 1989;264:1837-45.
37. Smith FJ, Irvine AD, Terron-Kwiatkowski A, Sandilands A, Campbell LE, Zhao Y, et al. Loss-of-function mutations in the gene encoding filaggrin cause ichthyosis vulgaris. Nat Genet 2006;38:337-42.
38. Novak N, Baurecht H, Schafer T, Rodriguez E, Wagenpfeil S, Klopp N, et al. Loss-of-function mutations in the filaggrin gene and allergic contact sensitization to nickel. J Invest Dermatol 2008;128:1430-5.
39. Weidinger S, Illig T, Baurecht H, Irvine AD, Rodriguez E, Diaz-Lacava A, et al. Loss-of-function variations within the filaggrin gene predispose for atopic dermatitis with allergic sensitizations. J Allergy Clin Immunol 2006;118:214-9.
40. Sergeant A, Campbell LE, Hull PR, Porter M, Palmer CN, Smith FJ, et al. Heterozygous null alleles in filaggrin contribute to clinical dry skin in young adults and the elderly. J Invest Dermatol 2009;129:1042-5.
41. Brown SJ, Relton CL, Liao H, Zhao Y, Sandilands A, McLean WH, et al. Filaggrin haploinsufficiency is highly penetrant and is associated with increased severity of eczema: further delineation of the skin phenotype in a prospective epidemiological study of 792 school children. Br J Dermatol 2009;161:884-9.
42. Jakasa I, Verberk MM, Bunge AL, Kruse J, Kezic S. Increased permeability for polyethylene glycols through skin compromised by sodium lauryl sulphate. Exp Dermatol 2006;15:801-7.
43. Jakasa I, Verberk MM, Esposito M, Bos JD, Kezic S. Altered penetration of polyethylene glycols into uninvolved skin of atopic dermatitis patients. J Invest Dermatol 2007;127:129-34.
44. Hata M, Tokura Y, Takigawa M, Sato M, Shioya Y, Fujikura Y, et al. Assessment of epidermal barrier function by photoacoustic spectrometry in relation to its importance in the pathogenesis of atopic dermatitis. Lab Invest 2002;82:1451-61.
45. Seidenari S, Giusti G. Objective assessment of the skin of children affected by atopic dermatitis: a study of pH, capacitance and TEWL in eczematous and clinically uninvolved skin. Acta Derm Venereol 1995;75:429-33.
46. Elias PM, Hatano Y, Williams ML. Basis for the barrier abnormality in atopic dermatitis: outside-inside-outside pathogenic mechanisms. J Allergy Clin Immunol 2008;121:1337-43.
47. Eberlein-Konig B, Schafer T, Huss-Marp J, Darsow U, Mohrenschlager M, Herbert O, et al. Skin surface pH, stratum corneum hydration, trans-epidermal water loss and skin roughness related to atopic eczema and skin dryness in a population of primary school children. Acta Derm Venereol 2000;80:188-91.
48. Nemoto-Hasebe I, Akiyama M, Nomura T, Sandilands A, McLean WH, Shimizu H. Clinical severity correlates with impaired barrier in filaggrin-related eczema. J Invest Dermatol 2009;129:682-9.
49. Hubiche T, Ged C, Benard A, Leaute-Labreze C, McElreavey K, de Verneuil H, et al. Analysis of SPINK 5, KLK 7 and FLG genotypes in a French atopic dermatitis cohort. Acta Derm Venereol 2007;87:499-505.

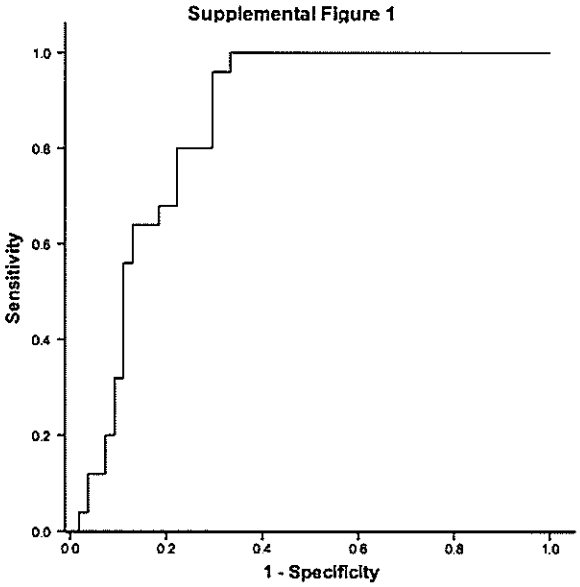



**FIG E1.** ROC curve for NMF values to distinguish $FLG^{+/-}$ subjects from $FLG^{-/-}$ subjects in the $AD_{FLG}$ group. AUC is 0.85 (95% CI, 0.76-0.93), with a sensitivity of 96% and a specificity of 66.77% at a cutoff of 0.767 arbitrary units.

Medical genetics

# A rare connexin 26 mutation in a patient with a *forme fruste* of keratitis–ichthyosis–deafness (KID) syndrome

Ching Yin Neoh, MBBS, MRCP, MMed(Int Med), Huijia Chen*, BSc(Hons), See Ket Ng, MBBS, MMed(Int Med), Ellen Birgitte Lane*, PhD, FRSE, and John Edmund Armourer Common*, PhD

From the National Skin Center and *Agency for Science, Technology and Research, Institute of Medical Biology, Singapore

**Correspondence**
Ching Yin Neoh, MBBS, MRCP, MMED(INT MED)
1 Mandalay Road
Singapore 308205
E-mail: mailcyneoh@yahoo.com.sg

## Abstract

**Background** Keratitis–ichthyosis–deafness (KID) syndrome is a rare ectodermal dysplasia characterized by generalized erythrokeratotic plaques, sensorineural hearing loss, and vascularizing keratitis. Cutaneous changes and hearing loss typically present in early childhood, whereas ocular symptoms present later. Mutations in the connexin (Cx) 26 gene, *GJB2*, are now established to underlie many of the affected cases, with the majority of patients harboring the p.D50N mutation.

**Methods** A rare patient demonstrating features of incomplete KID syndrome associated with an uncommon Cx26 gene mutation is described.

**Results** The patient presented late in adolescence with partial features of KID syndrome. There was limited cutaneous involvement and the rare association of cystic acne. Both hearing impairment and ophthalmic involvement were mild in severity. Genetic mutation analysis revealed a previously described, rare mutation in *GJB2*, resulting in a glycine to arginine change at codon 12 (p.G12R).

**Conclusions** This report describes a patient exhibiting characteristics suggestive of a late-onset, incomplete form of KID syndrome with the *GJB2* mutation (p.G12R). The p.G12R mutation has only been described in one other patient with KID syndrome, whose clinical presentation was not characterized.

## Introduction

The original features of keratitis–ichthyosis–deafness (KID) [Online Mendelian Inheritance in Man (OMIM): 148210] syndrome, described by Burns in 1915,[1] include congenital generalized erythroderma with deafness and keratitis. Since then, reports of KID syndrome with varying degrees of severity and involvement have surfaced.[1–3] Most cases of KID syndrome described to date are sporadic, although autosomal dominant and recessive patterns of inheritance have been reported.[1–3] The unifying genetic mutation in these patients is in the *GJB2* gene encoding connexin (Cx) 26, discovered in 2002.[4] The majority of cases have been associated with p.D50N, with further mutations associated with p.G12R, p.N14Y, p.S17F, p.D50Y, p.G45E and p.A40V.[4–8] p.D50N has also been associated with hystrix-like–ichthyosis–deafness (HID) (OMIM: 602540) syndrome and KID and HID syndromes have been proposed to represent a single disease entity.[9] Pathogenic mutations in *GJB6* (Cx30) also give rise to symptoms that are similar to KID syndrome, indicating genetic heterogeneity.[10] In this report, a patient demonstrating an incomplete or incipient variant of KID syndrome with a rare *GJB2* mutation is presented.

## Case Report

A 21-year-old Singaporean Chinese man, the younger of two children of a nonconsanguineous marriage, presented with cracked areas on his palms and soles and severe acne vulgaris present since the age of 15 years. His acne vulgaris was recalcitrant and only responded transiently to a course of isotretinoin. Bilateral mild sensorineural hearing loss was detected on routine school screening at 7 years of age. No hearing aids were needed. The patient had a long-standing history of "sensitive eyes" with left eye epithelial scarring and vascularization on ophthalmic examination. This had been treated symptomatically as keratoconjunctivitis by the ophthalmologist. His vision was otherwise good without correction. Developmental milestones were normal and no relevant family history was noted.



*International Journal of Dermatology* 2009, 48, 1078–1081

**Figure 1** Grainy, spiculated appearance of facial skin

**Figure 2** Plantar keratoderma with characteristic stippled and dotted appearance

Physical examination showed a well-developed, intellectually normal young man with coarse facial features, characterized by lichenification of the skin with deep furrows on the forehead. The skin over the temples had a grainy, spiculated appearance (Fig. 1). Multiple acne nodules and cysts were seen. The hair and dentition were normal. The palms and soles were hyperkeratotic, with a characteristic stippled and dotted appearance (Fig. 2). His fingernails were dystrophic with transverse ridges seen in a few. Some of his toenails were atrophic and brittle. Ophthalmologic examination revealed mild congestion of both conjunctivae. The patient also had recurrent tinea cruris as well as tinea pedis, which necessitated intermittent courses of oral and topical antifungal therapy.

Other differential diagnoses, such as Vohwinkel's syndrome, Bart–Pumphrey's syndrome, and Clouston's syndrome, were discounted because of a lack of additional

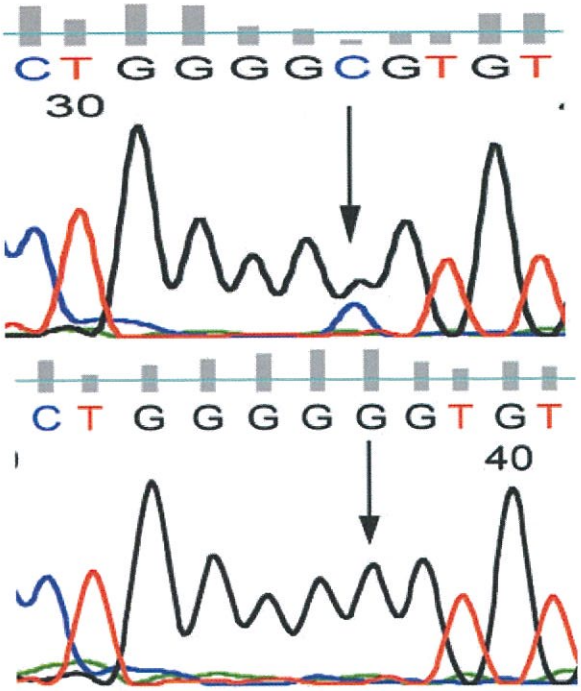



**Figure 3** Mutational analysis showing transversion 34G > C in *GJB2* in our patient (top sequence) and normal control (bottom sequence)

characteristic cutaneous findings and as a result of the ophthalmic involvement which is unique to KID syndrome.

Genomic DNA was extracted from the patient's peripheral blood sample following informed consent. Polymerase chain reaction (PCR) was completed with primers encapsulating the coding exons of *GJB2* (primer sequences available on request). Mutation analysis of *GJB2* was carried out. The PCR products were run on a 1% agarose gel to confirm fragment size before purification and direct sequencing using BigDye V3.1, according to the manufacturer's instructions (Applied Biosystems, California, USA). The patient was shown to harbor a transversion c.34G > C in *GJB2*, leading to the replacement of glycine with arginine (p.G12R) (Fig. 3). *GJB6* that encodes Cx30 was also fully analyzed and no genetic mutation was detected.

## Discussion

KID syndrome is a rare genetic connexin disorder with disturbed gap junction functions. Gap junctions are important in cellular communication and are formed by two unrelated protein families: the connexins and pannexins.[11] KID syndrome is manifested clinically as generalized erythrokeratoderma, profound sensorineural hearing loss, and vascularizing keratitis. A variety of genetic mutations on Cx26 and Cx30

have been described in association with the condition.[4–8,10] Although it is known that connexins play a key role in tissue homeostasis, the exact role of gap junction communication during keratinocyte differentiation has not been elucidated.[12] There is much clinical heterogeneity in KID syndrome. In the largest reported series of KID syndrome to date, patients varied in their presenting skin abnormalities, severity of ocular symptoms, as well as onset of hearing impairment.[12] In this report, we describe a patient in whom the clinical characteristics of KID syndrome were incompletely expressed. Nevertheless, he was found to harbor a Cx26 mutation associated with KID syndrome.

In contrast with the commonly described generalized thickening of skin seen in KID syndrome, our patient only had limited cutaneous involvement in the form of follicular hyperkeratotic lesions on the palms, soles, and face.[1–6,12] The late onset of skin involvement at 15 years of age was also atypical, as skin changes are usually present at birth or noted within the first few weeks of life.[3] Hearing loss and eye symptoms in our patient were hardly symptomatic and he did not require any special interventions for daily living. Early-onset profound sensorineural hearing loss and severe ocular complications encountered in other reported patients with KID syndrome were absent in our patient.[3,12–14] Ectodermal abnormalities, such as those involving the hair and dentition, were absent.[3]

Isolated nodulocystic acne vulgaris occurring in KID syndrome was another rare clinical feature demonstrated by our patient. Previous reports have described patients with nodulocystic acne as part of the follicular occlusion triad.[15,16] Mazereeuw-Hautier *et al.*[12] recently described the rare occurrence of inflammatory nodules in 43% of their patients. Interestingly, the histology of some of these nodules revealed *in situ* squamous cell carcinoma. Severe cystic acne also constitutes one of the features highlighted in two young patients with KID syndrome, who developed malignant proliferating pilar tumors.[17] It has been suggested that the underlying genetic defect in KID syndrome results in impaired keratinization of the epidermis and hair follicles.[15] Cx26 has also been considered as a tumor suppressor gene that modulates development, growth, and differentiation in epithelial tissues.[18] Stringent surveillance for skin and oral cancers in patients with KID syndrome is imperative, especially for patients with clinical evidence of follicular occlusion.[17,19] Although our patient did not strictly satisfy the criteria for the follicular occlusion triad, he is under close observation for the development of skin tumors in view of his recalcitrant cystic acne and recurrent skin infections.

An attempt to correlate the *GJB2* genotype to phenotype in KID syndrome was made recently.[12] It appeared that certain mutations, such as p.S17F and p.G45E, showed more severe skin involvement and a poorer prognosis. The p.G12R mutation seen in our patient has only been reported previously in one other patient of European descent, with no precise clinical description.[4] From our single patient with this rare mutation, we speculate that patients harboring this mutation may have less extensive skin involvement, possibly of a later onset, and associated with cystic acne.

In conclusion, we have described a rare patient demonstrating partial characteristics of KID syndrome in association with nodulocystic acne and an uncommon Cx26 mutation. This case draws attention to the importance of recognizing the milder, atypical phenotypes of KID syndrome, which may be easily left undiagnosed without appropriate genetic studies. Early diagnosis will allow for genetic counseling as well as long-term surveillance for malignancy and infections encountered in KID syndrome.

## Acknowledgment

This work was funded at the Institute of Medical Biology by Agency for Science, Technology and Research (A*STAR), Singapore.

## References

1 Burns FS. A case of generalized keratoderma with unusual involvement of the eyes and nasal and buccal mucous membranes. *J Cutan Dis* 1915; 33: 255–360.
2 Traupe H. Not an ichthyosis at all: the keratitis, ichthyosis-like hyperkeratosis and deafness (KID) syndrome. In: Traupe H, ed. *The Ichthyoses. A Guide to Clinical Diagnosis, Genetic Counseling and Therapy*. Berlin: Springer-Verlag, 1989: 198–202.
3 Caceres-Rios H, Tamayo-Sanchez L, Duran-Mckinster C, *et al.* Keratitis, ichthyosis and deafness (KID syndrome): review of the literature and proposal of a new terminology. *Paediatr Dermatol* 1996; 13: 105–113.
4 Richard G, Rouan F, Willoughby CE, *et al.* Missense mutations in *GJB2* encoding connexin-26 cause the ectodermal dysplasia keratitis–ichthyosis–deafness syndrome. *Am J Hum Genet* 2002; 70: 1341–1348.
5 Richard G. Connexin gene pathology. *Exp Dermatol* 2003; 28: 397–409.
6 Richard G. Connexin disorders of the skin. *Clin Dermatol* 2005; 23: 23–32.
7 Yotsumoto S, Hashiguchi T, Chen X, *et al.* Novel mutations in GJB2 encoding connexin-26 in Japanese patients with keratitis–ichthyosis–deafness syndrome. *Br J Dermatol* 2003; 148: 649–653.
8 Janecke AR, Hennies HC, Günther B, *et al.* GJB2 mutations in keratitis–ichthyosis–deafness syndrome including its fatal form. *Am J Med Genet* 2005; 133: 128–131.
9 van Geel M, van Steensel MA, Kuster W, *et al.* HID and KID syndromes are associated with the same connexin 26 mutation. *Br J Dermatol* 2002; 146: 938–942.
10 Jan AY, Amin S, Ratajczak P, *et al.* Genetic heterogeneity of KID syndrome: identification of a Cx30 gene (GJB6)

mutation in a patient with KID syndrome and congenital atrichia. *J Invest Dermatol* 2004; 122: 1108–1113.
11 Meşe G, Richard G, White TW. Gap junctions: basic structure and function. *J Invest Dermatol* 2007; 127: 2516–2524.
12 Mazereeuw-Hautier J, Bitoun E, Chevrant-Breton J, *et al.* Keratitis–ichthyosis–deafness syndrome: disease expression and spectrum of connexin 26 (*GJB2*) mutations in 14 patients. *Br J Dermatol* 2007; 156: 1015–1019.
13 Szymko-Bennett Y, Russell LJ, Bale SJ, *et al.* Auditory manifestations of keratitis–ichthyosis–deafness (KID) syndrome. *Laryngoscope* 2002; 112: 272–280.
14 Messmer EM, Kenyon KR, Rittinger O, *et al.* Ocular manifestations of keratitis–ichthyosis–deafness (KID) syndrome. *Ophthalmology* 2005; 112: 1–6.
15 Montgomery JR, White T, Martin B, *et al.* A novel connexin 26 gene mutation associated with features of the keratitis–ichthyosis–deafness syndrome and the follicular occlusion triad. *J Am Acad Dermatol* 2004; 51: 377–382.
16 Maintz L, Betz RC, Allam JP, *et al.* Keratitis–ichthyosis–deafness syndrome in association with follicular occlusion triad. *Eur J Dermatol* 2005; 15: 347–352.
17 Nyquist GG, Mumm C, Grau R, *et al.* Malignant proliferating pilar tumors arising in KID syndrome: a report of two patients. *Am J Med Genet* 2007; 143A: 734–741.
18 Lee SW, Tomasetto C, Sager R. Positive selection of candidate tumour-suppressor genes by subtractive hybridization. *Proc Natl Acad Sci USA* 1991; 88: 2825–2829.
19 Grob JJ, Breton A, Bonafe JL, *et al.* Keratitis, ichthyosis and deafness (KID) syndrome: vertical transmission and death from multiple squamous cell carcinomas. *Arch Dermatol* 1987; 123: 777–782.

# Association of Skin Barrier Genes within the PSORS4 Locus Is Enriched in Singaporean Chinese with Early-Onset Psoriasis

Huijia Chen[1,2], Terry K. L. Toh[3], Ildiko Szeverenyi[2], Rick T. H. Ong[3], Colin T. S. Theng[4], W. H. Irwin McLean[2], Mark Seielstad[3] and E. Birgitte Lane[1]

Psoriasis (OMIM#177900) is a common polygenic skin disorder affecting approximately 2% of the northern European population and 0.1% of the Han Chinese. Psoriasis patients suffer from chronic skin inflammation, manifested by erythematous scaly lesions. PSORS1–PSORS9 have been confirmed as psoriasis susceptibility loci in independent genetic studies on predominantly Caucasian populations, with psoriasis susceptibility loci (PSORS1, PSORS9) and additional loci at 9q33-34 and 2p22.3-11.2 reported in Han Chinese patients. In this study, we show the association of PSORS4 with psoriasis in Singaporean Chinese. Dense genotyping of single-nucleotide polymorphism-tagging candidate genes within the epidermal differentiation complex revealed significant association in the proximity of the involucrin gene (*IVL*); the strongest association was seen in early-onset psoriasis patients ($P=0.0014$). A follow-up genome-wide association screen localized the psoriasis susceptibility region to ~360 kb along chromosome 1 in the vicinity of *IVL*, small proline-rich region (*SPRR*) and proline-rich region 9 (*PRR9*) genes. The study of interactions between the causative variant(s) in this locus will provide insights into a possible role for epidermal barrier formation in the pathogenesis of psoriasis.

*Journal of Investigative Dermatology* (2009) **129,** 606–614; doi:10.1038/jid.2008.273; published online 11 September 2008

## INTRODUCTION

Psoriasis (OMIM#177900) is a common skin disorder affecting approximately 2% of the northern European population (Nevitt and Hutchinson, 1996) and 0.1% of the Han Chinese (Shao, 1996). Affected people suffer from chronic skin inflammation, manifested by erythematous scaly lesions that can be distributed over the scalp, palms, soles, and extensor surfaces of the limbs (Schon and Boehncke, 2005). At the microscopic level, psoriatic skin demonstrates features of abnormal differentiation (hyperkeratosis, parakeratosis) in the upper epidermal layers (Leigh *et al.*, 1985) and inflammation (neutrophil infiltration of the epidermis, T-cell infiltration of the dermis and dermal vessel dilatation; Schon and Boehncke, 2005).

The etiology of psoriasis is incompletely understood (Barker, 1998), with its manifestations resulting from the complex interplay of the immune system, the epidermal barrier, and environmental factors. Although streptococcal infection, medications, and trauma are involved in triggering psoriasis (Krueger and Duvic, 1994; Fry and Baker, 2007), greater attention has recently been focused on elucidating genetic factors that influence susceptibility to psoriasis (Bowcock and Cookson, 2004; Bowcock, 2005). The epidemiology of psoriasis is consistent with it being a polygenic disease where multiple genes with small effects interact with one another and with the environment to determine an individual's propensity to develop psoriasis (Bhalerao and Bowcock, 1998).

Independent studies of the major histocompatibility complex in different ethnic groups consistently show an association between psoriasis and the HLA-Cw6 allele (Tiilikainen *et al.*, 1980; Nair *et al.*, 1997; Trembath *et al.*, 1997; Enlund *et al.*, 1999a; Zhang *et al.*, 2002; Sagoo *et al.*, 2004; Lesueur *et al.*, 2007a). This locus at chromosome 6p21 has been dubbed psoriasis susceptibility locus 1 (PSORS1; Elder *et al.*, 1994a, b; Trembath *et al.*, 1997). Although PSORS1 is a predominant locus, it seems to be involved in only 30–50% of psoriasis cases (Asumalahti *et al.*, 2003). Linkage studies on small groups of individuals have identified other PSORS loci, including PSORS2 at 17q25 (Tomfohrde *et al.*, 1994), PSORS3 at 4qter (Matthews *et al.*, 1996), PSORS4 at 1q21 (Capon *et al.*, 1999), PSORS5 at 3q21 (Enlund *et al.*, 1999b), PSORS6 at 19p13 (Lee *et al.*, 2000), PSORS7 at 1p (Veal *et al.*, 2001), PSORS8 at 16q (Nair *et al.*, 1997), and PSORS9 at 4q31-34. Additional susceptibility loci and genes have also been postulated (International Psoriasis

[1]*A*STAR Institute of Medical Biology, Singapore, Singapore;* [2]*Epithelial Genetics Group, Division of Molecular Medicine, Colleges of Life Sciences and Medicine, Dentistry and Nursing, Medical Sciences Institute, Dundee, UK;* [3]*A*STAR Genome Institute of Singapore, Singapore, Singapore and* [4]*National Skin Centre, Singapore, Singapore*

*Correspondence: Professor E. Birgitte Lane, A*STAR Institute of Medical Biology, 8A Biopolis Grove, No. 06-06 Immunos, Singapore 138665, Singapore. E-mail: birgit.lane@imb.a-star.edu.sg*

*Abbreviations: EDC, epidermal differentiation complex; PSORS, psoriasis susceptibility locus; SNP, single-nucleotide polymorphism*

*Received 9 April 2008; revised 2 July 2008; accepted 15 July 2008; published online 11 September 2008*

Genetics Consortium, 2003; Capon *et al.*, 2004, 2008; Lesueur *et al.*, 2007b; Liu *et al.*, 2008). The results of these studies have proved difficult to replicate, because geographic variations in psoriasis prevalence are compatible with both genetic explanations that emphasize interethnic differences and environmental variations. For example, only the PSORS1, PSORS9 (Zhang *et al.*, 2002; Fan *et al.*, 2008), 9q33-34 (Sun *et al.*, 2007), and 2p22.3-11.2 (Sun *et al.*, 2008) regions have been reported as susceptibility loci in the Han Chinese. Moreover, the existence of more than one major susceptibility gene is likely to decrease the ability to detect linkage (Capon *et al.*, 2000).

PSORS4, localized within chromosome 1q21, first emerged as a potential PSORS from American and Italian family studies (Bhalerao and Bowcock, 1998; Capon *et al.*, 1999). Fine mapping with densely spaced markers in 22 families of Italian descent further narrowed the region to a 350 kb interval in the distal portion of the epidermal differentiation complex (EDC) (Capon *et al.*, 2001). Recently, a genome-wide association study of psoriasis and psoriatic arthritis patients in the United States and United Kingdom reported association of the disorder with a single-nucleotide polymorphism (SNP) in the late cornified envelope 1C (*LCE1C*) protein within the EDC (Liu *et al.*, 2008). The EDC spans 2 mb in 1q21 and contains at least 45 genes involved in skin cornification (Mischke *et al.*, 1996; Hoffjan and Stemmler, 2007), a differentiation process that leads to formation of a functional epidermal barrier. The product of these genes can be grouped into three families: the cornified envelope precursor proteins, the keratin filament-associated proteins and the S100 calcium-binding proteins. The first family includes loricrin (*LOR*), involucrin (*IVL*), small proline-rich proteins (*SPRR*s), and the family of late cornified envelope (*LCE*) proteins. During cornification, these proteins are cross-linked by transglutaminases to form a strong, insoluble cornified envelope. The second family consists of filaggrin (*FLG*), trichohyalin (*TCHH*), filaggrin2 (*FLG2*), repetin (*RPTN*; Krieg *et al.*, 1997), hornerin (*HRNR*; Makino *et al.*, 2001), and cornulin (*CRNN*; Contzler *et al.*, 2005). These proteins are important in bundling keratins to attain flattened squames during cornification. In addition to their structural role, it has been postulated that many of these genes also participate in calcium-dependent processes because they possess two EF-hand type calcium-binding domains (Huber *et al.*, 2005). The last family of genes in the EDC belong to the S100 family that include small, calcium-binding proteins involved in signaling and cell cycle progression (Kligman and Hilt, 1988; Eckert *et al.*, 2004).

As psoriasis displays abnormal epidermal differentiation, the genes in the EDC cluster are biologically plausible candidates for psoriasis susceptibility. In psoriatic patients, upregulation of *S100A7*, *S100A8*, and *S100A9* have been observed in the epidermis (Saintigny *et al.*, 1992; Hoffmann *et al.*, 1994) and *LOR* is downregulated (Giardina *et al.*, 2004). *HRNR* was reported to be expressed only in psoriatic and regenerating skin (Takaishi *et al.*, 2005).

The aim of this study was to identify gene(s) within the EDC that confer susceptibility to psoriasis in Singaporean Chinese psoriatic patients by high-density SNP genotyping within the 2 mb EDC region. The results provide evidence for an association of psoriasis within the vicinity of *IVL*, especially in early-onset psoriasis.

## RESULTS

### Clinical features

The majority of the psoriatics in the study had common plaque psoriasis (471); there were 7 with palmoplantar psoriasis, 9 with pustular psoriasis, and 4 with erythrodermic psoriasis. There were three individuals with overlapping features (one instance each of plaque and palmoplantar, plaque and pustular, plaque and erythrodermic). A total of 494 Singaporean Chinese psoriatic individuals were eventually studied; 360 had an age of onset $\leqslant$40 years whereas 129 were diagnosed after the age of 40 (there was no clinical data available for 5 individuals). For clarity purposes, all genotyping data presented is derived from patients diagnosed with common plaque psoriasis (474).

### SNP genotyping of EDC candidate genes on the Sequenom platform

The EDC locus maps to a 2.0 mb interval in chromosome 1q21 and contains at least 45 genes whose products are involved in the process of skin cornification (Mischke *et al.*, 1996; Hoffjan and Stemmler, 2007) We selected 11 candidate genes (Figure 1) from this region which was previously associated with psoriasis susceptibility (Capon *et al.*, 1999; Sun *et al.*, 2006). After SNP selection and screening, we had an initial set of 94 SNPs extending from positions 150346613 to 151500919 on chromosome 1 (NCBI 36). Once the analysis of this set was completed, a second set of 22 SNPs, clustered in a region that showed statistically significant association with psoriasis, was chosen from the Genome Institute of Singapore SNP database and genotyped together with 9 SNPs from the first set. Ultimately, 116 SNPs were genotyped (Table 1).

After discarding 22 SNPs whose genotype data did not conform to the Hardy–Weinberg equilibrium, the remaining 94 SNPs (listed in Table S1) were analysed. Significant association ($P<0.05$) was detected with 3 SNPs mapped in regions flanking *IVL* (Figure 1b; Table 2). Further statistical analysis on early-onset ($\leqslant$40 years old at diagnosis) plaque psoriasis patients revealed significant association with 11 SNPs clustered within 139 kb of the *IVL* region (Figure 1c). The lowest *P*-value (0.00143) was obtained with rs6661932 located 12.3 kb downstream of *IVL* (Table 2).

### Expanded analysis of EDC genes with the Infinium platform

The survey was expanded to the entire EDC with additional DNA samples; data from 267 SNPs distributed in a 2 mb interval from positions 150000000 to 152000000 of chromosome 1 (NCBI 36) were analyzed in 428 plaque psoriasis patients versus 571 normal controls after filtering out ethnic outliers and members of first degree relative pairs (Table S2). Statistically significant association ($P<0.05$) was still obtained with a SNP (rs2879485) mapped to the vicinity of the *IVL* gene. The expanded approach yielded an additional

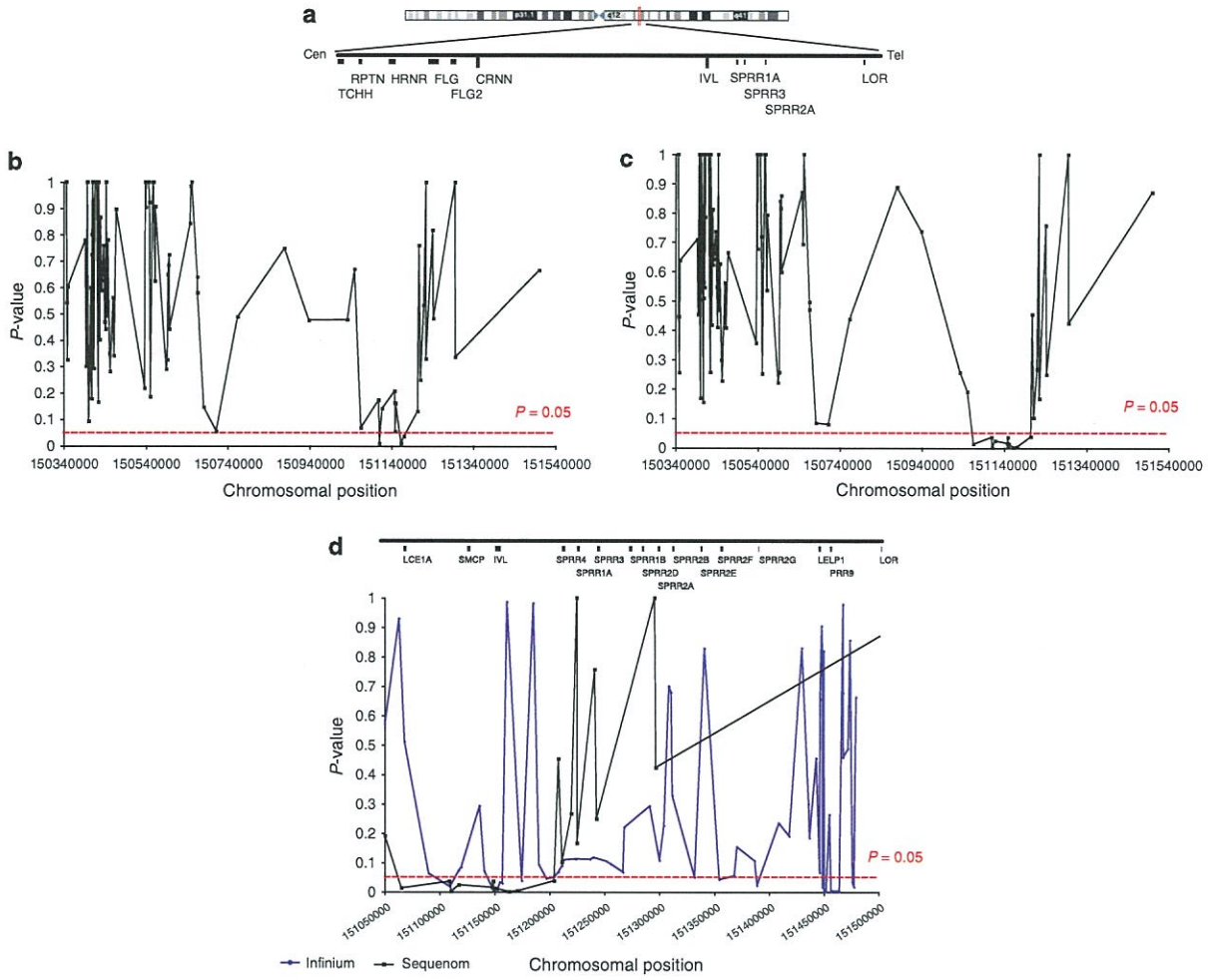



**Figure 1. SNP-based association analysis of candidate genes within chromosome 1q21.** (**a**) Schematic diagram of candidate genes chosen for Sequenom genotyping. (**b**) *P*-value plot of 94 SNPs genotyped by Sequenom in all plaque psoriasis patients; distribution of all these SNPs conforms to the Hardy–Weinberg equilibrium. The red dotted line indicates the threshold of statistical significance ($P<0.05$). (**c**) *P*-value plot of 94 SNPs genotyped by Sequenom in early-onset ($\leqslant$40 years old at diagnosis) plaque psoriasis patients; distribution of all these SNPs conforms to the Hardy–Weinberg equilibrium. The red dotted line indicates the threshold of statistical significance ($P<0.05$). (**d**) A high resolution map of the region flanked by late cornified envelope 1A gene and loricrin genes and the corresponding *P*-value plot of SNPs genotyped in early-onset ($\leqslant$40 years old) plaque psoriasis patients with Infinium (blue) and Sequenom (black) platforms showing the close proximity of possible PS susceptibility locus.

cluster of SNPs with statistical evidence for association with psoriasis (Table 3). These SNPs are further downstream of *IVL* and positioned within a 120 kb segment of EDC, between the *SPRR2A* and *LOR* genes, which were not previously selected for analysis with the Sequenom platform (Table 1). One of these SNPs was located in the intron of *SPRR2C* pseudogene, the rest were found in the neighborhood of the *SPRR2F*, *SPRR2G*, late cornified envelope-like proline-rich protein 1 (*LELP1*), and proline-rich protein 9 (*PRR9*) genes (Table 3). Concurrence of association in the above region was observed in the subanalysis of early-onset plaque psoriasis patients (Figure 1d); all SNPs located in *IVL* also showed significant association (Table S2).

The linkage disequilibrium plot for a subset of the EDC SNPs that encompass all that are statistically associated with psoriasis is displayed in Figure 2. rs2879485 forms a small linkage disequilibrium block of 5 kb (block 25) with rs1854779 (350 bp upstream of *IVL*). The remaining SNPs associated with psoriasis separate into 3 well-defined linkage disequilibrium blocks (blocks 32, 35, and 36 that span 34, 17, and 11 kb, respectively).

## DISCUSSION

Psoriasis has been perceived and studied primarily as a chronic inflammatory disorder (Bowcock, 2005). Changes in the amounts of inflammatory cells and cytokines in the dermis are thought to lead to epidermal hyperkeratosis and parakeratosis as a secondary phenomenon (Ghoreschi *et al.*, 2007; Nickoloff *et al.*, 2007). Recent advances in our understanding of atopic dermatitis (eczema) may necessitate a revision of this model. Recently, *FLG* has been shown to be the major gene responsible for ichthyosis vulgaris

**Table 1. List of candidate genes in the epidermal differentiation complex chosen for SNP genotyping**

| Candidate genes/ genomic regions | Chromosomal start | Chromosomal end | Length (kb) | No. of SNPs genotyped |
|---|---|---|---|---|
| *TCHH* | 150345417 | 150353180 | 7.76 | 5 |
| *RPTN* | 150392695 | 150398328 | 5.63 | 3 |
| Region between *RPTN* and *HRNR* | 150398329 | 150451175 | 52.8 | 35 |
| *HRNR* | 150451176 | 150463296 | 12.1 | 10 |
| *FLG* | 150541275 | 150564303 | 27.9 | 16 |
| *FLG2* | 150587837 | 150599106 | 11.3 | 8 |
| *CRNN* | 150648343 | 150653363 | 5.02 | 3 |
| Region between *CRNN* and *IVL* | 150653364 | 151147644 | 494 | 17 |
| *IVL* | 151147663 | 151150985 | 3.32 | 4 |
| Region between *IVL* and *SPRR1A* | 151150986 | 151223180 | 72.2 | 8 |
| *SPRR1A* | 151223181 | 151224913 | 1.73 | 2 |
| *SPRR3* | 151240849 | 151242956 | 2.11 | 2 |
| *SPRR2A* | 151295220 | 151296612 | 1.39 | 2 |
| *LOR* | 151498803 | 151501224 | 2.42 | 1 |
| | | | Total | 116 |

SNP, single-nucleotide polymorphism.

**Table 2. SNPs showing significant *P*-values in Sequenom genotyping**

| SNP | Chromosomal location | Gene annotation | Remarks | Major/minor allele | *P*-value (all cases versus controls) | *P*-value (early-onset cases versus controls) |
|---|---|---|---|---|---|---|
| rs4845325 | 151065184 | *LCE1A* intron | NA | G/A | 0.0686 | 0.0135 |
| rs2339396 | 151108924 | — | NA | G/A | 0.1737 | 0.0362 |
| *rs11205128 | 151110766 | — | NA | G/A | 0.0091 | 0.0016 |
| rs3737861 | 151117562 | *SMCP* intron | NA | T/G | 0.1414 | 0.0244 |
| rs16834751 | 151147835 | *IVL* intron | NA | A/C | 0.2079 | 0.0160 |
| rs11205132 | 151148759 | *IVL* intron | NA | A/G | 0.1628 | 0.0360 |
| rs2229496 | 151149234 | *IVL* exon 2 | Nonsynonymous | G/A | 0.0558 | 0.0051 |
| rs913996 | 151150735 | *IVL* 3′ UTR | NA | C/T | 0.1618 | 0.0129 |
| *rs6661932 | 151163358 | downstream of *IVL* | 12.3 kb | C/T | 0.0106 | 0.0014 |
| *rs4845497 | 151170786 | downstream of *IVL* | 19.8 kb | C/T | 0.0359 | 0.0037 |
| rs4845501 | 151204025 | — | NA | A/G | 0.1312 | 0.0379 |

NA, not applicable; SNP, single-nucleotide polymorphism.
The chromosomal location is based on NCBI build 36 and dbSNP 126.
*SNPs with significant *P*-values ($<0.05$) in all psoriasis cases versus controls and early-onset ($\leqslant 40$ years old at diagnosis) psoriasis cases versus controls.

(Smith *et al.*, 2006) and also contributes significantly to the onset of atopic dermatitis and asthma in atopic dermatitis patients (Palmer *et al.*, 2006). Loss-of-function *FLG* mutations were observed in as many as 50% of moderate-severe atopic dermatitis cases (Sandilands *et al.*, 2007) and also confer major susceptibility to early-onset atopic dermatitis (Barker *et al.*, 2007). It has been suggested that the abnormal cornification resulting from the *FLG* mutations disrupts the epidermal barrier, allowing percutaneous entry of antigens that trigger the inflammation observed in atopic dermatitis. A similar pathogenetic mechanism in which inflammation occurs secondarily to epidermal disarray (Albanesi *et al.*, 2007) may explain the association of psoriasis with skin trauma. The search for psoriasis susceptibility genes should therefore be widened to include genomic regions that are important in skin cornification (Tschachler, 2007).

The proteins expressed by the genes encoded by the EDC, of which filaggrin is one, are crucial in cornification. The EDC itself maps to chromosome 1q21, the same locus as PSORS4, which has previously been reported as a PSORS in Europeans (Capon *et al.*, 1999) and Americans (Bhalerao

**Table 3. Human HapMap 550v3 SNPs along the EDC within chromosome 1q21 showing significant *P*-values in Infinium genotyping**

| SNP | Chromosomal location | Gene annotation | Remarks | Major/minor allele | *P*-value (all cases versus controls |
|---|---|---|---|---|---|
| rs1011297 | 150864572 | Upstream of *LCE3A* | 2639 bp | T/C | 0.0378 |
| rs7543194 | 150940844 | Downstream of *LCE2A* | 2303 bp | T/C | 0.0477 |
| rs2879485 | 151152537 | Downstream of *IVL* | 1551 bp | G/A | 0.0294 |
| rs1500941 | 151354285 | Downstream of *SPRR2F* | 1672 bp | A/G | 0.0212 |
| rs428913 | 151368112 | Upstream of *SPRR2G* | 11479 bp | C/T | 0.0319 |
| rs509194 | 151388795 | *SPRR2C*[1] intron | NA | A/G | 0.0134 |
| rs10494291 | 151448915 | Downstream of *LELP1* | 4695 bp | G/T | 0.0375 |
| rs10494292 | 151450469 | Downstream of LELP1/ upstream of PRR9 | 6249 bp/6215 bp | T/G | 0.0012 |
| rs10788861 | 151452018 | Upstream of *PRR9* | 4666 bp | C/A | 0.0023 |
| rs1410859 | 151456561 | Upstream of *PRR9*[2] | 123 bp | G/A | 0.0121 |
| rs1410860 | 151456602 | Upstream of *PRR9*[2] | 82 bp | A/G | 0.0033 |
| rs12127862 | 151464064 | Downstream of *PRR9* | 5647 bp | G/A | 0.0031 |
| rs10888541 | 151476212 | Downstream of *PRR9* | 17795 bp | A/G | 0.0345 |
| rs4845342 | 151477432 | Downstream of *PRR9* | 19015 bp | C/A | 0.0353 |

EDC, epidermal differentiation complex; NA, not applicable; SNP, single-nucleotide polymorphism.
[1]Unprocessed pseudogene.
[2]Possible promoter region of *PRR9*.

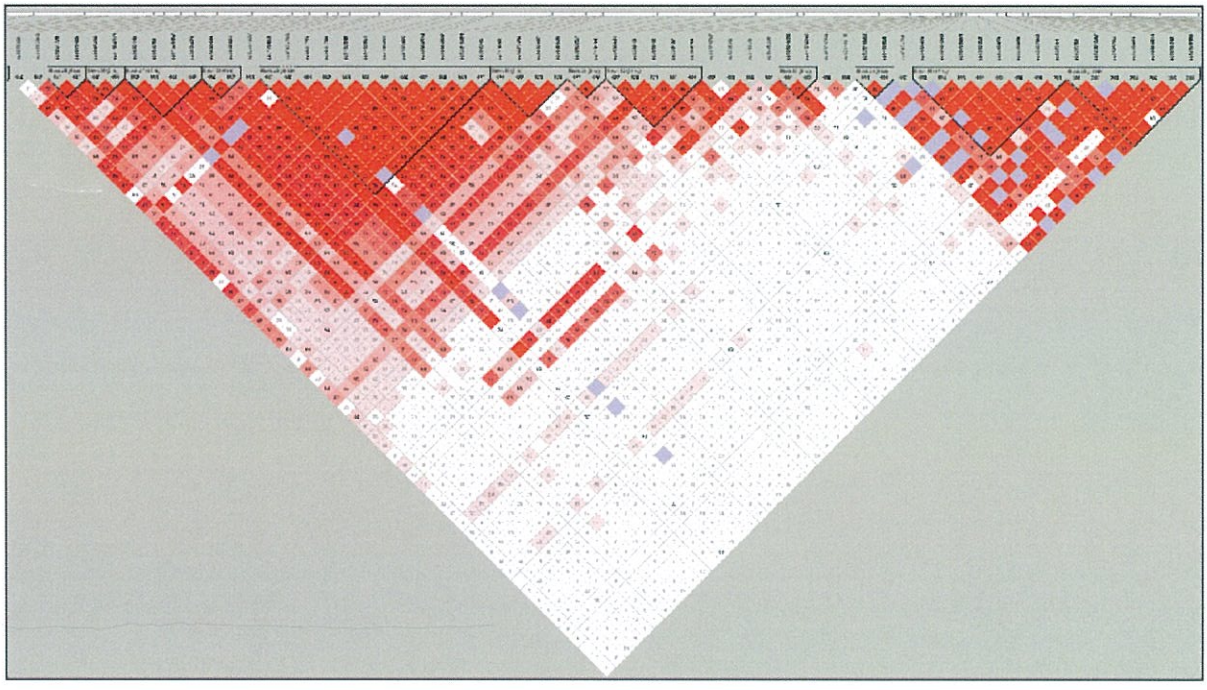


**Figure 2. Linkage disequilibrium (LD) plot of 62 contiguous SNPs from the Human HapMap 550v3 panel, including all that are significantly associated with psoriasis.** This subset of EDC SNPs extends from positions 151136387 to 151479292 (dbSNP 126) and includes all that are significantly associated with psoriasis. They group into 12 haplotype blocks (blocks 25–36). SNPs that are excluded from the blocks were filtered using criteria specified in the text.

and Bowcock, 1998). However, this locus has not been previously associated with psoriasis in any Asian population. Involucrin, filaggrin, and loricrin are the major components of the cornified envelope. Previous analyses of mutations and SNPs in the *FLG* and *LOR* genes indicate that both of them are unlikely to be the gene(s) that confer susceptibility to

psoriasis in PSORS4 (Giardina *et al.*, 2004; Zhao *et al.*, 2007). However, SNPs from the *LCE* family cluster and peptidoglycan recognition protein coding region, that lie between *FLG* and *S100A9* genes, have shown association with psoriasis in a genome-wide association study and family-based Caucasian patient cohort respectively (Sun *et al.*, 2006; Liu *et al.*, 2008).

Our preliminary association study of 11 selected genes within the EDC revealed 3 SNPs in the vicinity of *IVL* with significant *P*-values ($0.0090 < P < 0.040$; Table 2); subanalysis of early-onset patients revealed significant association of 11 SNPs ($0.0016 < P < 0.040$) clustered within 139 kb of *IVL*, highlighting the presence of a strong genetic component in early-onset plaque psoriasis (Henseler and Christophers, 1985). Involucrin was the first structural protein to be identified as a cornified envelope component and has been reported to be show abnormal expression in psoriatic skin (Ishida-Yamamoto and Iizuka, 1995; Pena-Penabad *et al.*, 1999). Involucrin is an early component in the process of skin cornification—it becomes cross-linked to many structural proteins in the epidermal layer, serving as a scaffold to aid the subsequent maturation of the stratum corneum. Regulation of involucrin expression is complex and involves spatially distinct regulatory elements (Crish *et al.*, 2002). The proximal regulatory region containing AP1-1 and C/EBP transcription factor binding sites lies directly upstream of exon 1 whereas the distal regulatory region between −2473/−1953 include ISE, AP1-5, and Sp1 binding sites (Crish *et al.*, 2006). Of the 11 SNPs, 4 that scored as significant in early-onset psoriasis are within the *IVL* gene, 4 others are in close proximity to the distal regulatory region, and two others are 12.3 and 19.8 kb downstream of *IVL*, respectively (Table 2). This region may also harbor transcription factor binding sites that participate in the complex regulation of the involucrin gene previously alluded to.

Our subsequent expanded analysis of the EDC revealed further interesting insights. Borderline evidence ($0.030 < P < 0.050$) was detected with 2 SNPs in the *LCE* family and a SNP in the flanking region of *IVL*; another distinct cluster of 11 SNPs significantly associated with psoriasis ($0.0014 < P < 0.049$) was identified within a 120 kb region of EDC 200 kb distal from *IVL* (Table 3; Figure 1c). This region includes *SPRR2F SPRR2G, SPRR2C, PRR9,* and *LELP1* genes. A literature search revealed that these genes have been less intensively studied than other EDC components. SPRRs, also known as cornifins, are precursors of the cornified envelope and are more conspicuously expressed in psoriatic skin (Iizuka *et al.*, 2004). *LELP1* has not previously been associated with psoriasis, although SNPs in its proximity are associated with elevated immunoglobulin E levels, a feature of atopic dermatitis (Sharma *et al.*, 2007). The majority of the SNPs are positioned in the surrounding region of *PRR9*, of which rs14100859 and rs1410860 are polymorphisms that could potentially affect promoter activity (Table 3). Interestingly, subanalysis of early-onset plaque psoriasis patients revealed significant association ($P < 0.05$) for all *IVL* SNPs, supporting our Sequenom genotyping data and reaffirming the possible involvement of this region in psoriasis susceptibility (Table S2).

Our data confirming PSORS4 as a susceptibility locus in Singaporean Chinese are consistent with previous findings in Caucasians that polymorphisms in precursor proteins of cornified envelope genes are associated with susceptibility to psoriasis (Capon *et al.*, 2001; Sun *et al.*, 2006; Liu *et al.*, 2008). The observation of enhanced association of PSORS4 in early-onset psoriasis patients supports the hypothesis that an altered skin barrier from birth can expose an individual to more environmental irritants, leading to an early-onset of inflammatory skin disorders like psoriasis and atopic dermatitis (Barker *et al.*, 2007). Despite the recent discovery of *FLG* as an important gene in the maintenance of an efficient epidermal barrier (Irvine and McLean, 2006), our data confirms, in line with Zhao *et al.* (2007), that *FLG* is probably not important in psoriasis (Figure 1b).

In conclusion, we have identified two clusters of SNPs lying among genes responsible for skin barrier formation in the PSORS4 locus that show significant association with psoriasis in Singaporean Chinese. These clusters are enriched in the early-onset patients. It will be interesting to see if our findings are replicated in other populations, especially in neighboring Asian regions.

## MATERIALS AND METHODS

### Clinical samples

Singaporean Chinese patients with psoriasis were clinically ascertained and recruited through the Psoriasis Clinic of the National Skin Centre, Singapore, and unaffected Singaporean Chinese with a similar age and gender distribution to the patients were selected as controls from stored samples in the Singapore Tissue Network. Informed consent was obtained under protocols approved by the Institutional Review Boards of the participating institutions, which are in accordance with the Declaration of Helsinki Principles.

DNA extraction was performed within the Singapore Tissue Network facilities using the Gentra PureGene Cell Kit (Qiagen, Hilden, Germany; www.qiagen.com).

The first analysis was performed with DNA extracted from 405 patients and 431 controls, whereas the second experiment was carried out with the same set of samples and additional DNA, collected from a total of 494 patients and 588 controls.

### Candidate gene and SNP selection

For Sequenom analysis, 11 candidate genes were chosen by virtue of their function in skin differentiation and localization to the distal 350 kb of the EDC (Table 1). All SNPs within these candidates and in 100 kb regions upstream and downstream of them were identified through the Genome Institute of Singapore SNP collective database. A subset of biallelic SNPs spaced at approximately 1.5 kb intervals in the region of these candidate genes was selected, with priority given to those which had previously been validated and shown to be present in the Han Chinese with minor allele frequency $\geqslant 10\%$ wherever possible. These were then screened with the RealSNP tool (Sequenom, San Diego, CA; www.realsnp.com) to remove SNPs with ambiguous or highly repetitive flanking sequences.

In order to further analyze the initial findings, we decided to examine the entire EDC with 267 SNPs from the Human HapMap 550v3 panel (Illumina, San Diego, CA) that were located in an interval spanning positions 150000000 to 152000000 in

chromosome 1 based on NCBI build 36 and dbSNP build 126 data (Table S2).

### Genotyping

Two sets of assay were performed using different genotyping technologies. SNPs selected from the Genome Institute of Singapore SNP database were genotyped on 405 patients and 431 controls using Sequenom MassARRAY technology (Sequenom; www.sequenom.com). The iPLEX assay was performed according to the manufacturer's instructions (www.sequenom.com). Briefly, PCR and extension reactions were designed by the MassARRAY design software, and were carried out using 5 ng of template DNA. The unincorporated nucleotides in the PCR product were deactivated with shrimp alkaline phosphatase. SNP sites were amplified using MassExtend primers in a termination master mix with differentially labeled ddNTPs. The primer extension products were desalted with the Clean Resin, spotted onto a SpectroCHIP bioarray, and MALDI-TOF MS analysis was performed with MassARRAY Workstation version 3.3 software. Resulting spectra were analyzed and genotypes were determined using the Sequenom TYPER software.

The Human HapMap SNPs were genotyped with the Infinium assay protocol (Illumina; www.illumina.com) on 750 ng of genomic DNA each from 494 patients and 588 controls (including all individuals from the first set) using the standard reagent kit provided. Genotypes were assigned with BeadStudio 3.0 software (Illumina).

### Statistical analysis

Genotype data derived from the Sequenom platform was analyzed with HelixTree 2.0 software (Golden Helix, Bozeman, MT; www.goldenhelix.com). Allele-based association studies were performed with missing alleles excluded from analysis. Analysis for conformation with the Hardy–Weinberg equilibrium was performed with a $\chi^2$-test to identify potential errors in genotype calls, and SNPs that yielded *P*-values $<0.05$ were discarded.

The genotype data from the Infinium platform was initially subjected to a quality check with BeadStudio 3.0 and SNPs with call frequencies $<88\%$ and samples with call rates $<96\%$ were discarded. Data from the remaining 469 patients and 585 controls was analyzed with PLINK, a toolset for whole genome association analysis (http://pngu.mgh.harvard.edu/~purcell/plink). Three filters were applied to the data:

(i) a gender concordance check;
(ii) detection of ethnic outliers using principal components analysis as implemented in the program EIGENSTRAT (Price *et al.*, 2006);
(iii) detection of first-degree relatives (one of each pair was dropped from subsequent analysis for association).

The Cochran–Armitage TREND test was then used for association analysis on the filtered data whereas conformation with the Hardy–Weinberg equilibrium was checked using the exact test (Wigginton *et al.*, 2005). Age-stratified analysis of the filtered data was performed with 313 early-onset ($\leqslant$40 years old at diagnosis) and 111 late-onset ($>$40 years old at diagnosis) plaque psoriasis patients after discarding data with no information on the age of onset.

The Infinium genotype data from the controls was analyzed for linkage disequilibrium using Haploview 4.0 (http://www.broad.mit.edu/mpg/haploview). Standard filtering procedures were implemented (Hardy–Weinberg, *P*-value $<0.001$, percentage of non-missing genotypes $<75\%$, number of Mendelian errors $>0$, minor allele frequency $<0.001$).

## CONFLICT OF INTEREST

The authors declare that they have no conflict of interest.

## ACKNOWLEDGMENTS

We thank the psoriasis patients at the National Skin Centre, Singapore for volunteering their DNA, which made this research possible. We also thank Carine Bonnard, Chan Kian Mun, Jason Ong, and Meah Wee Yang for genotyping assistance and Dr Lawrence Khoo Shih Wee of the National Skin Centre. Research in the Institute of Medical Biology and Genome Institute of Singapore is funded by the Biomedical Research Council of the Agency for Science, Technology and Research (A*STAR), Singapore. Skin barrier research in the McLean laboratory is supported by grants from the British Skin Foundation, the National Eczema Society, the Medical Research Council (Reference number G0700314) and anonymous donations from the Tayside region of Scotland.

## SUPPLEMENTARY MATERIAL

**Table S1.** Complete list of SNPs that satisfy the Hardy–Weinberg equilibrium (HWE) along the EDC and their *P*-values in Sequenom genotyping.

**Table S2.** Complete list of SNPs along the EDC and their *P*-values in Infinium genotyping.

## REFERENCES

Albanesi C, De Pita O, Girolomoni G (2007) Resident skin cells in psoriasis: a special look at the pathogenetic functions of keratinocytes. *Clin Dermatol* 25:581–8

Asumalahti K, Laitinen T, Lahermo P, Suomela S, Itkonen-Vatjus R, Jansen C *et al.* (2003) Psoriasis susceptibility locus on 18p revealed by genome scan in Finnish families not associated with PSORS1. *J Invest Dermatol* 121:735–40

Barker JN (1998) Pathogenesis of psoriasis. *J Dermatol* 25:778–81

Barker JN, Palmer CN, Zhao Y, Liao H, Hull PR, Lee SP *et al.* (2007) Null mutations in the filaggrin gene (FLG) determine major susceptibility to early-onset atopic dermatitis that persists into adulthood. *J Invest Dermatol* 127:564–7

Bhalerao J, Bowcock AM (1998) The genetics of psoriasis: a complex disorder of the skin and immune system. *Hum Mol Genet* 7:1537–45

Bowcock AM (2005) The genetics of psoriasis and autoimmunity. *Annu Rev Genomics Hum Genet* 6:93–122

Bowcock AM, Cookson WO (2004) The genetics of psoriasis, psoriatic arthritis and atopic dermatitis. *Hum Mol Genet* 13 Spec No 1 R43–55

Capon F, Bijlmakers MJ, Wolf N, Quaranta M, Huffmeier U, Allen M *et al.* (2008) Identification of ZNF313/RNF114 as a novel psoriasis susceptibility gene. *Hum Mol Genet* 17:1938–45

Capon F, Dallapiccola B, Novelli G (2000) Advances in the search for psoriasis susceptibility genes. *Mol Genet Metab* 71:250–5

Capon F, Novelli G, Semprini S, Clementi M, Nudo M, Vultaggio P *et al.* (1999) Searching for psoriasis susceptibility genes in Italy: genome scan and evidence for a new locus on chromosome 1. *J Invest Dermatol* 112:32–5

Capon F, Semprini S, Chimenti S, Fabrizi G, Zambruno G, Murgia S *et al.* (2001) Fine mapping of the PSORS4 psoriasis susceptibility region on chromosome 1q21. *J Invest Dermatol* 116:728–30

Capon F, Trembath RC, Barker JN (2004) An update on the genetics of psoriasis. *Dermatol Clin* 22:339–47, vii

Contzler R, Favre B, Huber M, Hohl D (2005) Cornulin, a new member of the "fused gene" family, is expressed during epidermal differentiation. *J Invest Dermatol* 124:990–7

Crish JF, Bone F, Banks EB, Eckert RL (2002) The human involucrin gene contains spatially distinct regulatory elements that regulate expression during early versus late epidermal differentiation. *Oncogene* 21:738–47

Crish JF, Gopalakrishnan R, Bone F, Gilliam AC, Eckert RL (2006) The distal and proximal regulatory regions of the involucrin gene promoter have distinct functions and are required for *in vivo* involucrin expression. *J Invest Dermatol* 126:305–14

Eckert RL, Broome AM, Ruse M, Robinson N, Ryan D, Lee K (2004) S100 proteins in the epidermis. *J Invest Dermatol* 123:23–33

Elder JT, Nair RP, Guo SW, Henseler T, Christophers E, Voorhees JJ (1994a) The genetics of psoriasis. *Arch Dermatol* 130:216–24

Elder JT, Nair RP, Voorhees JJ (1994b) Epidemiology and the genetics of psoriasis. *J Invest Dermatol* 102:24S–7S

Enlund F, Samuelsson L, Enerback C, Inerot A, Wahlstrom J, Yhr M *et al.* (1999a) Analysis of three suggested psoriasis susceptibility loci in a large Swedish set of families: confirmation of linkage to chromosome 6p (HLA region), and to 17q, but not to 4q. *Hum Hered* 49:2–8

Enlund F, Samuelsson L, Enerback C, Inerot A, Wahlstrom J, Yhr M *et al.* (1999b) Psoriasis susceptibility locus in chromosome region 3q21 identified in patients from southwest Sweden. *Eur J Hum Genet* 7:783–90

Fan X, Yang S, Huang W, Wang ZM, Sun LD, Liang YH *et al.* (2008) Fine mapping of the psoriasis susceptibility locus PSORS1 supports HLA-C as the susceptibility gene in the Han Chinese population. *PLoS Genet* 4:e1000038

Fry L, Baker BS (2007) Triggering psoriasis: the role of infections and medications. *Clin Dermatol* 25:606–15

Ghoreschi K, Weigert C, Rocken M (2007) Immunopathogenesis and role of T cells in psoriasis. *Clin Dermatol* 25:574–80

Giardina E, Capon F, De Rosa MC, Mango R, Zambruno G, Orecchia A *et al.* (2004) Characterization of the loricrin (LOR) gene as a positional candidate for the PSORS4 psoriasis susceptibility locus. *Ann Hum Genet* 68:639–45

Henseler T, Christophers E (1985) Psoriasis of early and late onset: characterization of two types of psoriasis vulgaris. *J Am Acad Dermatol* 13:450–6

Hoffjan S, Stemmler S (2007) On the role of the epidermal differentiation complex in ichthyosis vulgaris, atopic dermatitis and psoriasis. *Br J Dermatol* 157:441–9

Hoffmann HJ, Olsen E, Etzerodt M, Madsen P, Thogersen HC, Kruse T *et al.* (1994) Psoriasin binds calcium and is upregulated by calcium to levels that resemble those observed in normal skin. *J Invest Dermatol* 103:370–5

Huber M, Siegenthaler G, Mirancea N, Marenholz I, Nizetic D, Breitkreutz D *et al.* (2005) Isolation and characterization of human repetin, a member of the fused gene family of the epidermal differentiation complex. *J Invest Dermatol* 124:998–1007

Iizuka H, Takahashi H, Honma M, Ishida-Yamamoto A (2004) Unique keratinization process in psoriasis: late differentiation markers are abolished because of the premature cell death. *J Dermatol* 31:271–6

International Psoriasis Genetics Consortium (2003) The International Psoriasis Genetics Study: assessing linkage to 14 candidate susceptibility loci in a cohort of 942 affected sib pairs. *Am J Hum Genet* 73:430–7

Irvine A, McLean W (2006) Breaking the (un)sound barrier: filaggrin is a major gene for atopic dermatitis. *J Invest Dermatol* 126:1200–2

Ishida-Yamamoto A, Iizuka H (1995) Differences in involucrin immunolabeling within cornified cell envelopes in normal and psoriatic epidermis. *J Invest Dermatol* 104:391–5

Kligman D, Hilt DC (1988) The S100 protein family. *Trends Biochem Sci* 13:437–43

Krieg P, Schuppler M, Koesters R, Mincheva A, Lichter P, Marks F (1997) Repetin (Rptn), a new member of the "fused gene" subgroup within the S100 gene family encoding a murine epidermal differentiation protein. *Genomics* 43:339–48

Krueger GG, Duvic M (1994) Epidemiology of psoriasis: clinical issues. *J Invest Dermatol* 102:14S–8S

Lee YA, Ruschendorf F, Windemuth C, Schmitt-Egenolf M, Stadelmann A, Nurnberg G *et al.* (2000) Genomewide scan in german families reveals evidence for a novel psoriasis-susceptibility locus on chromosome 19p13. *Am J Hum Genet* 67:1020–4

Leigh IM, Pulford KA, Ramaekers FC, Lane EB (1985) Psoriasis: maintenance of an intact monolayer basal cell differentiation compartment in spite of hyperproliferation. *Br J Dermatol* 113:53–64

Lesueur F, Lefevre C, Has C, Guilloud-Bataille M, Oudot T, Mahe E *et al.* (2007a) Confirmation of psoriasis susceptibility loci on chromosome 6p21 and 20p13 in French families. *J Invest Dermatol* 127:1403–9

Lesueur F, Oudot T, Heath S, Foglio M, Lathrop M, Prud'homme J *et al.* (2007b) ADAM33, a new candidate for psoriasis susceptibility. *PLoS ONE* 2:e906

Liu Y, Helms C, Liao W, Zaba LC, Duan S, Gardner J *et al.* (2008) A genome-wide association study of psoriasis and psoriatic arthritis identifies new disease loci. *PLoS Genet* 4:e1000041

Makino T, Takaishi M, Morohashi M, Huh NH (2001) Hornerin, a novel profilaggrin-like protein and differentiation-specific marker isolated from mouse skin. *J Biol Chem* 276:47445–52

Matthews D, Fry L, Powles A, Weber J, McCarthy M, Fisher E *et al.* (1996) Evidence that a locus for familial psoriasis maps to chromosome 4q. *Nat Genet* 14:231–3

Mischke D, Korge BP, Marenholz I, Volz A, Ziegler A (1996) Genes encoding structural proteins of epidermal cornification and S100 calcium-binding proteins form a gene complex ("epidermal differentiation complex") on human chromosome 1q21. *J Invest Dermatol* 106:989–92

Nair RP, Henseler T, Jenisch S, Stuart P, Bichakjian CK, Lenk W *et al.* (1997) Evidence for two psoriasis susceptibility loci (HLA and 17q) and two novel candidate regions (16q and 20p) by genome-wide scan. *Hum Mol Genet* 6:1349–56

Nevitt GJ, Hutchinson PE (1996) Psoriasis in the community: prevalence, severity and patients' beliefs and attitudes towards the disease. *Br J Dermatol* 135:533–7

Nickoloff BJ, Xin H, Nestle FO, Qin JZ (2007) The cytokine and chemokine network in psoriasis. *Clin Dermatol* 25:568–73

Palmer CN, Irvine AD, Terron-Kwiatkowski A, Zhao Y, Liao H, Lee SP *et al.* (2006) Common loss-of-function variants of the epidermal barrier protein filaggrin are a major predisposing factor for atopic dermatitis. *Nat Genet* 38:441–6

Pena-Penabad C, de Unamuno P, Garcia Silva J, Ludena MD, Gonzalez Sarmiento R, Perez-Arellano JL (1999) Altered expression of immunoreactive involucrin in lamellar ichthyosis. *Eur J Dermatol* 9:197–201

Price AL, Patterson NJ, Plenge RM, Weinblatt ME, Shadick NA, Reich D (2006) Principal components analysis corrects for stratification in genome-wide association studies. *Nat Genet* 38:904–9

Sagoo GS, Tazi-Ahnini R, Barker JW, Elder JT, Nair RP, Samuelsson L *et al.* (2004) Meta-analysis of genome-wide studies of psoriasis susceptibility reveals linkage to chromosomes 6p21 and 4q28-q31 in Caucasian and Chinese Hans population. *J Invest Dermatol* 122:1401–5

Saintigny G, Schmidt R, Shroot B, Juhlin L, Reichert U, Michel S (1992) Differential expression of calgranulin A and B in various epithelial cell lines and reconstructed epidermis. *J Invest Dermatol* 99:639–44

Sandilands A, Terron-Kwiatkowski A, Hull P, O'regan G, Clayton T, Watson R *et al.* (2007) Comprehensive analysis of the gene encoding filaggrin uncovers prevalent and rare mutations in ichthyosis vulgaris and atopic eczema. *Nat Genet* 39:650–4

Schon MP, Boehncke WH (2005) Psoriasis. *N Engl J Med* 352:1899–912

Shao CG (1996) The prevalence, prevention and treatment of psoriasis in China. *Chin J Dermatol* 29:75–6

Sharma M, Mehla K, Batra J, Ghosh B (2007) Association of a chromosome 1q21 locus in close proximity to a late cornified envelope-like proline-rich 1 (LELP1) gene with total serum IgE levels. *J Hum Genet* 52:378–83

Smith FJ, Irvine AD, Terron-Kwiatkowski A, Sandilands A, Campbell LE, Zhao Y *et al.* (2006) Loss-of-function mutations in the gene encoding filaggrin cause ichthyosis vulgaris. *Nat Genet* 38:337–42

Sun C, Mathur P, Dupuis J, Tizard R, Ticho B, Crowell T *et al.* (2006) Peptidoglycan recognition proteins Pglyrp3 and Pglyrp4 are encoded from the epidermal differentiation complex and are candidate genes for the Psors4 locus on chromosome 1q21. *Hum Genet* 119:113–25

Sun LD, Li W, Yang S, Fan X, Yan KL, Liang YH *et al.* (2007) Evidence for a novel psoriasis susceptibility locus at 9q33-9q34 in Chinese Hans. *J Invest Dermatol* 127:1140–4

Sun LD, Yang S, Liu JJ, Ren YQ, Fan X, Xu SX *et al.* (2008) Follow-up analysis of 180 Chinese Han families: identification of a novel locus for psoriasis at 2p22.3-11.2. *Br J Dermatol* 158:512–7

Takaishi M, Makino T, Morohashi M, Huh NH (2005) Identification of human hornerin and its expression in regenerating and psoriatic skin. *J Biol Chem* 280:4696–703

Tiilikainen A, Lassus A, Karvonen J, Vartiainen P, Julin M (1980) Psoriasis and HLA-Cw6. *Br J Dermatol* 102:179–84

Tomfohrde J, Silverman A, Barnes R, Fernandez-Vina MA, Young M, Lory D *et al.* (1994) Gene for familial psoriasis susceptibility mapped to the distal end of human chromosome 17q. *Science* 264:1141–5

Trembath RC, Clough RL, Rosbotham JL, Jones AB, Camp RD, Frodsham A *et al.* (1997) Identification of a major susceptibility locus on chromosome 6p and evidence for further disease loci revealed by a two stage genome-wide search in psoriasis. *Hum Mol Genet* 6:813–20

Tschachler E (2007) Psoriasis: the epidermal component. *Clin Dermatol* 25:589–95

Veal CD, Clough RL, Barber RC, Mason S, Tillman D, Ferry B *et al.* (2001) Identification of a novel psoriasis susceptibility locus at 1p and evidence of epistasis between PSORS1 and candidate loci. *J Med Genet* 38:7–13

Wigginton JE, Cutler DJ, Abecasis GR (2005) A note on exact tests of Hardy-Weinberg equilibrium. *Am J Hum Genet* 76:887–93

Zhang X, He P, Wang Z, Zhang J, Li Y, Wang H *et al.* (2002) Evidence for a major psoriasis susceptibility locus at 6p21(PSORS1) and a novel candidate region at 4q31 by genome-wide scan in Chinese hans. *J Invest Dermatol* 119:1361–6

Zhao Y, Terron-Kwiatkowski A, Liao H, Lee SP, Allen MH, Hull PR *et al.* (2007) Filaggrin Null Alleles Are Not Associated with Psoriasis. *J Invest Dermatol* 127:1878–82

# Analysis of the individual and aggregate genetic contributions of previously identified serine peptidase inhibitor Kazal type 5 *(SPINK5)*, kallikrein-related peptidase 7 *(KLK7)*, and filaggrin *(FLG)* polymorphisms to eczema risk

**Stephan Weidinger, MD, PhD,[a,b]* Hansjörg Baurecht, MSc,[b,c]* Stefan Wagenpfeil, PhD,[c,d]* John Henderson, MD,[e] Natalija Novak, MD,[f] Aileen Sandilands, PhD,[g] Huijia Chen, BSc,[g] Elke Rodriguez, PhD,[b] Grainne M. O'Regan, MRCPI,[h] Rosemarie Watson, MD,[h] Haihui Liao, MD,[g] Yiwei Zhao, MD,[g] Jonathan N. W. N. Barker, MD,[i] Michael Allen, PhD,[i] Nick Reynolds, MD,[j] Simon Meggitt, MD,[j] Kate Northstone, PhD,[k] George D. Smith, DSc,[l] Carolin Strobl, MSc,[m] Caroline Stahl, MSc,[a,c] Thomas Kneib, PhD,[m] Norman Klopp, PhD,[n] Thomas Bieber, MD,[e] Heidrun Behrendt, MD,[b] Colin N. A. Palmer, PhD,[o] H.-Erich Wichmann, MD,[n] Johannes Ring, MD,[a,d] Thomas Illig, PhD,[n] W. H. Irwin McLean, PhD, DSc, FRSE,[g]* and Alan D. Irvine, MD[h,p]*** *Munich and Bonn, Germany, Bristol, Dundee, London, and Newcastle, United Kingdom and Dublin, Ireland*

**Background:** Polymorphisms in the serine protease inhibitor gene serine peptidase inhibitor Kazal type 5 *(SPINK5)* and the serine protease kallikrein-related peptidase 7 gene *(KLK7)* appear to confer risk to eczema in some cohorts, but these findings have not been widely replicated. These genes encode proteins thought to be involved in the regulation of posttranslation processing of filaggrin *(FLG)*, the strongest identified genetic risk factor for eczema to date.

**Objectives:** We sought to clarify the individual risk of eczema conferred by the *SPINK5* polymorphism rs2303067 (Glu420Lys) and a previously described insertion in the 3′ untranslated region of *KLK7* and to examine potential epistatic effects between these variants and *FLG* mutations.

**Methods:** Initially, we examined the effects of these polymorphisms and *FLG* in 486 unrelated patients from a German family-based study, an additional 287 German patients, and 418 unrelated Irish/English patients with eczema (n for 3 genes studied = 1191 vs 4544 control subjects). We then additionally studied the *SPINK5* polymorphism and *FLG* mutations in 1583 patients with eczema from the Avon Longitudinal Study of Parents and Children cohort (sample size for 2 genes studied = 2774 vs 10,607 control subjects).

**Results:** No association was seen with the *SPINK5* or *KLK7* variants in the case-control analysis; however, a weaker effect was observed for the *SPINK5* variant with maternal transmission in the family-based study. No interactions were seen between the polymorphisms in *KLK7*, *SPINK5*, and *FLG*.

**Conclusion:** The *SPINK5* 420LysSer mutation confers a risk of eczema when maternally inherited but is not a major eczema

From [a]the Department of Dermatology and Allergy Biederstein, Technical University Munich; [b]the Division of Environmental Dermatology and Allergy, Helmholtz Zentrum Munich–German Research Center for Environmental Health and ZAUM–Center for Allergy and Environment, Technical University Munich; [c]the Institute for Medical Statistics and Epidemiology IMSE, Technical University Munich; [d]the Graduate School of Information Science in Health (GSISH), Technical University Munich; [e]the Department of Community-Based Medicine, University of Bristol; [f]the Department of Dermatology and Allergy, University of Bonn; [g]the Epithelial Genetics Group, Human Genetics Unit, Division of Pathology and Neuroscience, University of Dundee; [h]the Department of Paediatric Dermatology, Our Lady's Children's Hospital, Crumlin, Dublin; [i]the St John's Institute of Dermatology, Guys Hospital, London; [j]the Department of Dermatology, University of Newcastle; [k]the Department of Social Medicine, University of Bristol; [l]MRC Centre for Causal Analyses in Translational Epidemiology, University of Bristol; [m]the Department for Statistics, Ludwig-Maximilans University, Munich; [n]the Department of Epidemiology, Helmholtz Zentrum Munich–German Research Center for Environmental Health; [o]the Population Pharmacogenetics Group, Biomedical Research Centre, University of Dundee; and [p]the Department of Clinical Medicine, Trinity College Dublin.

*These authors contributed equally to this work.

Supported by the German Ministry of Education and Research (BMBF) as part of the National Genome Research Network (NGFN) with grant NUW-S31T05. S. Weidinger and S. Wagenpfeil are supported by research grants KKF-07/04 and KKF-27/05 of the University Hospital "Rechts der Isar," Technical University Munich. In addition, Dr Weidinger is supported by a grant from the "Wilhelm-Vaillant Stiftung." The UK Medical Research Council, the Wellcome Trust, and the University of Bristol provide core support for Avon Longitudinal Study of Parents and Children. Additional funding was obtained from the Department of Paediatric Dermatology, Our Lady's Children's Hospital, Dublin, and by a grant from Tenovus (Tayside) to C.N.P. C.N.P. is supported by the Chief Scientists Office of the Scottish Executive Generation Scotland Initiative. The McLean laboratory is supported by grants from British Skin Foundation/National Eczema Association, the Pachyonychia Congenita Project, the Dystrophic Epidermolysis Bullosa Research Association, the Medical Research Council (reference no. G0700314), and anonymous donations from atopic families in the Tayside region of Scotland.

Disclosure of potential conflict of interest: N. Reynolds has served as an investigator in an investigator-initiated study for Stiefel, UK, and as an expert witness for Newcastle University on antipsoriatic therapy in research. J. Ring has received research support from Novartis, Schering-Plough, Fujisawa, GlaxoSmithKline, Bencard, Stallergenes, ALK-Abelló, Allergopharma, Pharmacia, DPC Biermann, Aventis, Almirall, Leo, Galderma, and Switch Biotech and has served as president of Allergopharma. A. D. Irvine has received funding from the Children's Medical and Research Foundation. The rest of the authors have declared that they have no conflict of interest.

Received for publication March 13, 2008; revised April 28, 2008; accepted for publication May 23, 2008.

Reprint requests: Stephan Weidinger, MD, PhD, Department of Dermatology and Allergy Biederstein, Technical University Munich (TUM), Munich 80802, Germany (or Alan D. Irvine, MD, Paediatric Dermatology, Our Lady's Children's Hospital, Crumlin, Dublin 12, Ireland). E-mail: weidinger@lrz.tum.de or irvinea@tcd.ie.

0091-6749/$34.00



doi:10.1016/j.jaci.2008.05.050

**risk factor. The *KLK7* insertion appears to confer no risk of eczema. We found no interaction between the *SPINK5* risk allele or the putative *KLK7* risk allele and *FLG* mutations. (J Allergy Clin Immunol 2008;122:560-8.)**

***Key words:*** *Eczema, atopy, skin barrier, stratum corneum, epistasis*

Eczema is a common chronic inflammatory skin disease with a complex, multifactorial cause and a strong genetic component.[1,2] Like other complex diseases, eczema is hypothesized to be determined by many genetic factors interacting with environmental components.[2] In the most commonly accepted paradigm for complex diseases, single genetic factors are considered to contribute only a modest amount to the total variation in the trait but are likely to exert additive or synergistic effects known as epistatic interactions.[3]

The identification of 2 common (R501X and 2282del4) and several rare mutations within the filaggrin gene *(FLG)* causing a deficiency of this key protein involved in skin barrier function delineated a major genetic risk for eczema.[4-6] Subsequently, an impressive series of replication studies[6-20] confirmed that these polymorphisms confer an exceptionally strong risk for eczema and subsequent allergen sensitization and that *FLG* is one of the strongest known genes for complex diseases in general.[21-24]

These observations suggest that the breakdown of the epidermal barrier represents one of the primary events in the development of eczema. This breach might then allow increased penetration of antigens, allergens, and irritants from the environment and thereby predispose to allergic sensitization and aberrant responses to microbial infection.[25-27]

Filaggrin is initially synthesized as biologically inactive profilaggrin, which is expressed as a highly phosphorylated insoluble protein in the granular layer of the epidermis. During the transition from granular cells to flattened squames, profilaggrin is processed to biologically active filaggrin monomers through several dephosphorylation and proteolytic steps,[28,29] the impairment of which might also impair skin barrier function. One of the proteases that has been suggested to be implicated in profilaggrin processing is the stratum corneum chymotryptic enzyme (SSCE),[30,31] which is possibly regulated by the serine protease inhibitor lymphoepithelial Kazal-type inhibitor (LEKTI), encoded by the serine peptidase inhibitor Kazal type 5 gene *(SPINK5)*.[31-33] Interestingly, an insertion in the 3′ untranslated region (UTR) of the kallikrein-related peptidase 7 gene *(KLK7)* encoding SCCE[34] has been reported to be associated with eczema. Early genome-wide linkage analysis of eczema family studies suggested a potential locus on 5q31, and after identification of 6 common polymorphisms in *SPINK5*, the variant Lys420Ser was associated with eczema in a cohort of British children.[35] This association has been replicated in 2 small Japanese studies,[36,37] but other studies have failed to replicate this association.

However, whereas *FLG* has been firmly established as a major gene for eczema, the reported effects of *KLK7* and *SPINK5* variants are rather weak and thus far lack robust confirmation in replication cohort studies. Therefore the aim of the present study was to address the existing literature and to clarify the role of these previously reported polymorphisms in *SPINK5* or *KLK7* in eczema. In addition, given their potential effects on posttranslational modification of filaggrin, we also sought to examine gene-gene interactions between *FLG*, *KLK7*, and *SPINK5*.

*Abbreviations used*
AIC: Akaike Information Criterion
ALSPAC: Avon Longitudinal Study of Parents and Children
*FLG*: Filaggrin gene
*KLK7*: Kallikrein-related peptidase 7 gene
KORA: Co-operative Health Research in the Region of Augsburg
LEKTI: Lymphoepithelial Kazal-type inhibitor
OR: Odds ratio
SNP: Single nucleotide polymorphism
*SPINK5*: Serine peptidase inhibitor, Kazal type 5 gene
SSCE: Stratum corneum chymotryptic enzyme
UK: United Kingdom
UTR: Untranslated region

## METHODS

### Study populations

*SPINK5*, *KLK7*, and *FLG* variants were typed in a cohort of 486 German parent–offspring trios for eczema, a collection of 418 English and Irish patients with eczema and 552 Irish blood donor control subjects, an additional series of 287 patients with eczema from Germany, and the population-based cross-sectional Co-operative Health Research in the Region of Augsburg (KORA) S4 cohort (n = 3992). In addition, the population-based Avon Longitudinal Study of Parents and Children (ALSPAC) cohort (n = 7646) was typed for the Glu420Lys *SPINK5* polymorphism and the 2 most common *FLG* mutations, R501X and 2282del4. Finally, to increase power, we performed a pooled analysis on all available data from all cohorts (for details on study populations, see Table E1 in this article's Online Repository at www.jacionline.org).

The study designs have been described in detail elsewhere.[9,18,38] Briefly, KORA S4 represents a sample of the general adult population of German nationality in the region of Augsburg recruited from October 1999 to April 2001. The survey comprised 4261 unrelated men and women between 25 and 74 years of age. All subjects had to complete a standardized questionnaire that, in addition to demographic data, included the basic allergy questions of the European Community Respiratory Health Survey on respiratory health.[39] All individuals received a skin examination by experienced senior dermatologists who had been additionally trained before the start of the study, according to the criteria of Hanifin and Rajka[40] and the United Kingdom (UK) diagnostic criteria for eczema.[41]

All German patients with eczema were unrelated and of white origin, with eczema diagnosed on the basis of a skin examination by experienced dermatologists using the UK diagnostic criteria.[41] In the family collection 10.4% of the parents had eczema (9.3% affected fathers and 11.1% affected mothers).

Patients with eczema from Ireland were recruited through attendance at a hospital-based clinic in Our Lady's Children's Hospital Crumlin, and the diagnosis was made according to the UK diagnostic guidelines by an experienced pediatric dermatologist (ADI, GO'R, or RW). The English eczema cohort was recruited from hospital-based clinics in London and Newcastle and has been described previously.[11] A summary of the demographics for all eczema study populations examined is presented in Table E1.

The ALSPAC is a longitudinal, population-based birth cohort study that recruited 14,541 unrelated pregnant women resident in Avon, UK, with expected dates of delivery between April 1, 1991, and December 31, 1992. There were 14,062 liveborn children. The study protocol has been described previously,[42,43] and further details are on the ALSPAC Web site (http://www.alspac.bris.ac.uk). At 6, 18, 30, and 42 months of age, the mothers were asked whether their children had skin rashes in the joints or creases of the body. As in previous studies, we defined individuals with eczema as those with reports of flexural dermatitis at 2 time points between 6 and 42 months.[25,42,43]

All study methods were approved by the relevant local authorities, and written informed consent that complies with all the Declaration of Helsinki Principles was obtained from all participants.

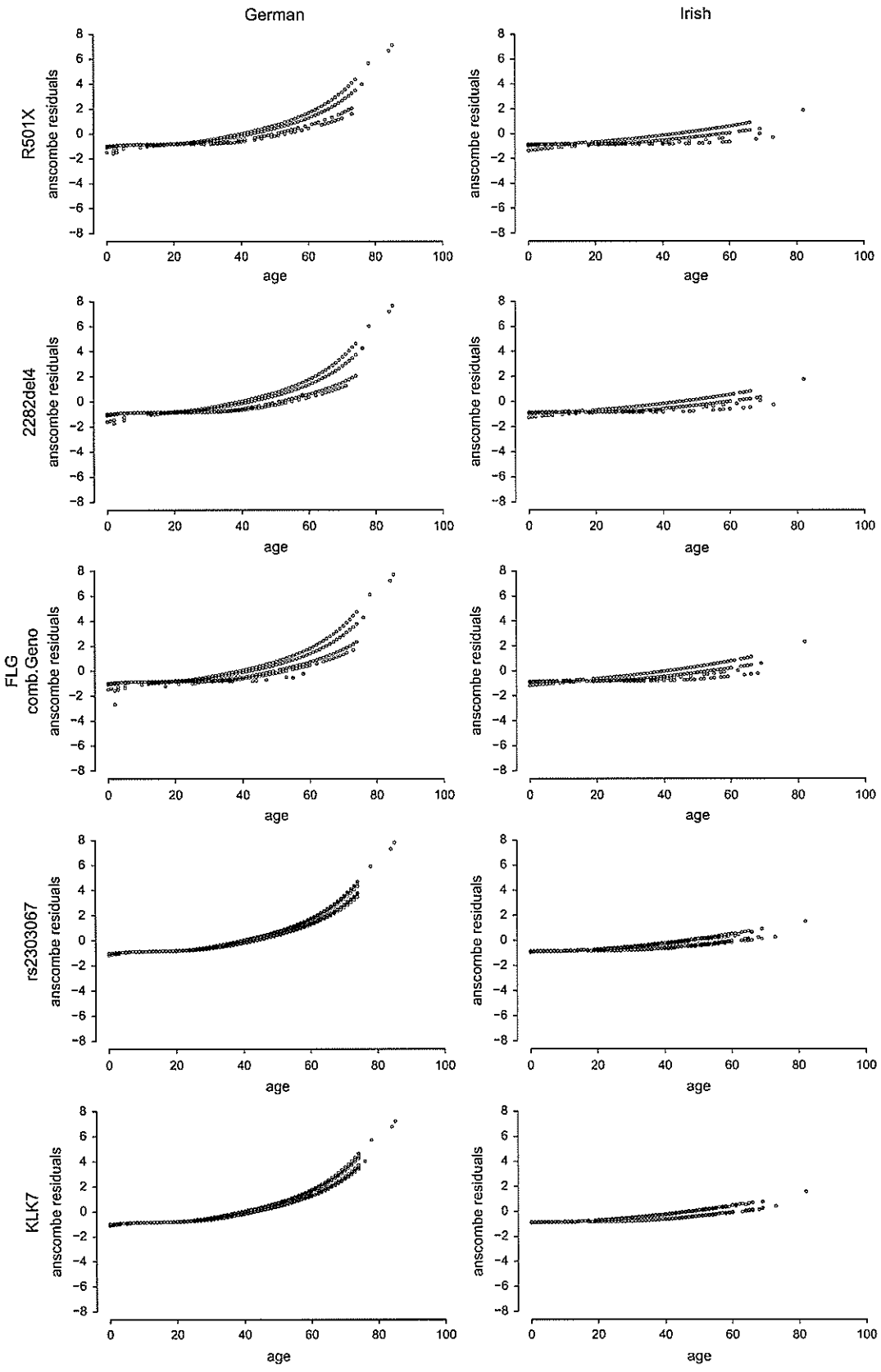



FIG 1. *(Continued).*

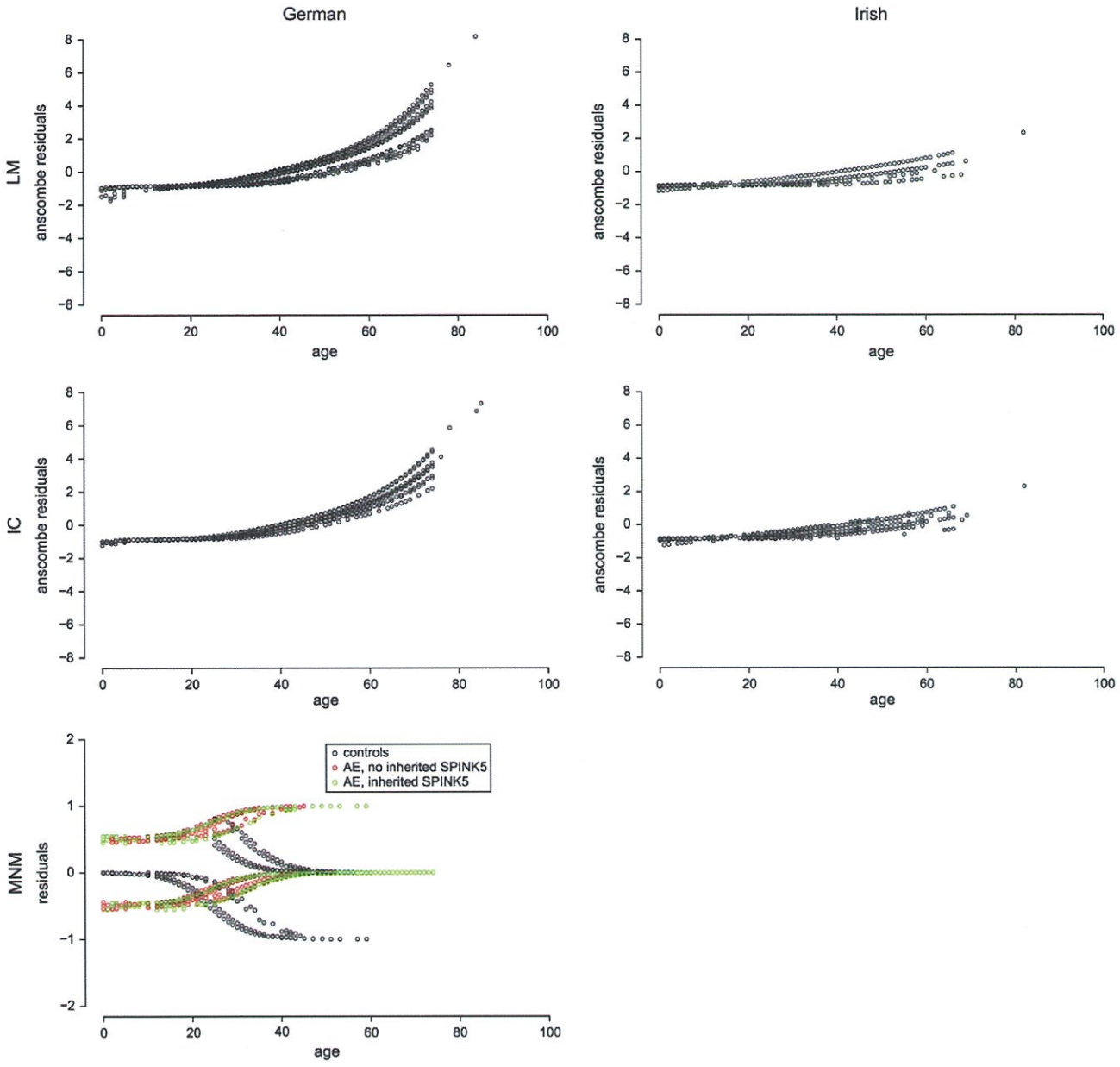



**FIG 2.** Residual plot of age after fitting a general linear model for SNP-SNP interaction analyses. The Anscombe residuals for LM and IC and the ordinary residuals for MNM show a functional form and do not spread randomly. Hence a general additive model is fitted to the data. The *columns* indicate the population's origin of the case-control studies, and the *rows* denote the applied methods. *LM,* Logistic regression model with product-interaction terms; *IC,* logistic regression model with interaction score; *MNM,* multinomial logistic model considering maternal inheritance.

## Genotyping

Genotyping in German samples was performed with the MassARRAY system (Sequenom, San Diego, Calif), as described previously.[9] Genotyping calls were made in real time with MASSARRAY RT software (Sequenom). Primers, as well as allele frequencies, in the population-based KORA S4 cohort (n = 4198) are shown Table E2 (available in this article's Online Repository at www.jacionline.org).

The Glu420Lys polymorphism in *SPINK5* (rs2303067) was typed in the Irish and English eczema cohorts and control subjects by using a predesigned single nucleotide polymorphism (SNP) Taqman Genotyping Assay from Applied Biosystems (Foster City, Calif; product C_2000249_10) and run on a 7900HT Fast Real-Time PCR system using the manufacturer's recommended protocol. The 4-bp insertion in the 3′UTR of *KLK7* was typed in Irish and English eczema cohorts and control subjects by sizing of fluorescently labeled

**FIG 1.** Residual plot of age after fitting a general linear model for individual SNP analyses. The Anscombe residuals show a functional form and do not spread randomly. Hence a general additive model is fitted for the data. comb.Geno, Combined genotype.

**TABLE I.** ORs and 95% CIs for associations between polymorphisms and eczema in the family-based analysis

| | Gene | Polymorphism | T:U | OR | 95% CI | | *P* value |
|---|---|---|---|---|---|---|---|
| German families (A) | *FLG* | r501x | 41:12 | 3.42 | 1.84 | 6.70 | $1.2 \times 10^{-4}$ |
| | | 2282del4 | 72:31 | 2.32 | 1.54 | 3.57 | $8.1 \times 10^{-5}$ |
| | | Combined genotype | 110:40 | 2.75 | 1.93 | 3.98 | $1.8 \times 10^{-8}$ |
| | *SPINK5* | rs2303067 | 259:207 | 1.25 | 1.04 | 1.50 | .01815 |
| | *KLK7* | AACC ins | 183:177 | 1.03 | 0.84 | 1.27 | .79215 |

*T*, Transmitted; *U*, untransmitted; *ins*, insertion.

**TABLE II.** Associations between polymorphisms and eczema in the case-control approaches

| | Gene | Polymorphism | OR | 95% CI | | *P*-value |
|---|---|---|---|---|---|---|
| German (A, B, E) | *FLG* | Combined genotype (het vs wt) | 4.15 | 3.10 | 5.56 | $1.3 \times 10^{-21}$ |
| Cases: 773 | | Combined genotype (hom vs wt) | 23.11 | 7.23 | 73.89 | $1.2 \times 10^{-7}$ |
| Control subjects: 3992 | *SPINK5* | rs2303067 (het vs wt) | 1.14 | 0.87 | 1.49 | .34285 |
| Total: 4765 | | rs2303067 (hom vs wt) | 1.22 | 0.89 | 1.67 | .21252 |
| | *KLK7* | AACC ins (het vs wt) | 1.11 | 0.89 | 1.40 | .35472 |
| | | AACC ins (hom vs wt) | 0.95 | 0.64 | 1.42 | .81033 |
| Irish/English (C, F) | *FLG* | Combined genotype (het vs wt) | 4.34 | 2.77 | 6.79 | $1.3 \times 10-10$ |
| Cases: 418 | | Combined genotype (hom vs wt) | $2.6 \times 10^{56}$ | 0 | ∞ | 1.0 |
| Control subjects: 552 | *SPINK5* | rs2303067 (het vs wt) | 0.78 | 0.52 | 1.18 | .23945 |
| Total: 970 | | rs2303067 (hom vs wt) | 1.15 | 0.71 | 1.87 | .57398 |
| | *KLK7* | (het vs wt) | 1.29 | 0.89 | 1.86 | .17756 |
| | | (hom vs wt) | 0.91 | 0.47 | 1.76 | .76988 |
| ALSPAC (D, G) | *FLG* | Combined genotype (het vs wt) | 2.17 | 1.81 | 2.60 | $3.3 \times 10^{6}$ |
| Cases: 1583 | | Combined genotype (hom vs wt) | 3.3 × 106 | 0 | ∞ | .94524 |
| Control subjects: 6063 | *SPINK5* | rs2303067 (het vs wt) | 0.96 | 0.84 | 1.09 | .52176 |
| Total: 7646 | | rs2303067 (hom vs wt) | 1.14 | 0.98 | 1.34 | .09156 |
| Pooled* (A, B, C, D, E, F, G) | *FLG* | Combined genotype (het vs wt) | 3.04 | 2.68 | 3.44 | $8.5 \times 10-68$ |
| Cases: 2774 | | Combined genotype (hom vs wt) | 49.38 | 19.72 | 123.61 | $8.2 \times 10-17$ |
| Control subjects: 10,607 | *SPINK5* | rs2303067 (het vs wt) | 0.97 | 0.87 | 1.07 | .50850 |
| Total: 13,381 | | rs2303067 (hom vs wt) | 1.13 | 1.00 | 1.27 | .04545 |

*het*, Heterozygous; *wt*, wild-type; *hom*, homozygous.
All logistic regression models are adjusted for age and sex. Every block refers to a single model in which only estimates of the genetic variables are displayed. No specific genetic model was assumed, and estimates are given for every genotype compared with the wild-type. Populations as designated in Table E1 are indicated in parentheses.
*For the pooled analyses, we estimated in all models a population effect as a confounder with *P* values of less than $10^{-4}$.

PCR products on an Applied Biosystems 3130xl Genetic Analyser. Ten-milliliter PCR reactions were performed with 25 ng of genomic DNA with 400 nmol/L forward primer (5′ gtt tct tca agt gtg caa gtt cac caa 3′) and 400 nmol/L FAM-labeled reverse primer (5′ GAT TGG TTT ATC AAC AGG GC 3′) in AmpliTaq Gold Buffer containing 1.5 mmol/L $MgCl_2$, 10 nmol of each deoxyribonucleoside triphosphate, 4% vol/vol dimethyl sulfoxide, and 0.25 U of AmpliTaq Gold polymerase (Applied Biosystems). PCR reactions were amplified at an annealing temperature of 58°C. Diluted PCR products were sized against ROX-500 size standards (Applied Biosystems). Allele sizes were 201 and 205 bp (insertion).

For the ALSPAC cohort, the 2 most common *FLG* mutations were typed as previously described,[24] and the *SPINK5* Glu420Lys polymorphism was typed by means of Taqman assay. Because we discovered 3 negative associations for *KLK7* in the eczema cohorts, we did not perform *KLK7* analysis in this very large population cohort.

## Statistical analyses

Descriptive statistics for quantitative and qualitative values are presented as means ± SD and relative frequencies or absolute numbers, respectively. Deviation from Hardy-Weinberg equilibrium was tested in parents for the family analyses and in control subjects for the case-control analyses. In the family setting we analyzed association of single SNPs with eczema using the classical transmission disequilibrium test. Parent-of-origin effects were investigated with the method proposed by Weinberg.[44]

Case-control analyses for single SNPs were performed with logistic regression models adjusted for age and sex. To not constrain the analyses to a specific genetic model, we modeled the 3 categorical genotypes by 2 dummy variables.

Gene-gene interaction analyses were performed after excluding individuals with 2 mutant *FLG* alleles because these individuals do not express the filaggrin protein, and therefore a biologic interaction with posttranslational modifications by SSCE and LEKTI is not plausible.

Four different approaches were carried out to estimate interaction effects between the single polymorphisms in *FLG*, *KLK7*, and *SPINK5*. In any of these approaches, we adjusted the models for the common covariates age and sex.

First, interaction was evaluated by using the logistic regression model with product-interaction terms. The Akaike Information Criterion (AIC) was used to select the appropriate model.[45]

Second, we defined an interaction score that counts the number of copies of the potentially disease-associated alleles,[46] which we used as a covariate in the logistic model.

Third, we modeled maternal effects observed for *SPINK5* using affected offspring from families only and estimated odds ratios (ORs) for 2 different affection statuses compared with control subjects in a multinomial regression model, which was carried out with BayesX 1.50.[47] An elaborate description of the statistical methods is given in this article's Methods section of the Online Repository (available at www.jacionline.org). For any of these regression approaches, the quantitative covariate age was modeled nonparametrically in a general additive model framework because residuals indicate structure in general linear models[48] (see Figs 1 and 2).

Finally, for further exploration, variable importance measures were computed by means of the random forest method. Random forests provide variable importance measures that can be used to detect variables relevant for

**TABLE III.** Associations between polymorphisms or gene-gene interaction terms and eczema

| | | | | | | |
|---|---|---|---|---|---|---|
| **German** | | | | | | |
| LRM (A, B, E) | **Gene-gene interaction term** | | **OR** | **95% CI** | | ***P* value** |
| | *FLG* combined genotype (het vs wt) | | 6.11 | 4.09 | 9.12 | $3.3 \times 10^{-10}$ |
| | rs2303067 (het vs wt) | | 1.28 | 0.87 | 1.90 | .11862 |
| | rs2303067 (hom vs wt) | | 1.28 | 0.83 | 1.97 | .18638 |
| | *FLG* combined genotype (het) rs2303067 (het) | | 0.62 | 0.34 | 1.11 | .17805 |
| | *FLG* combined genotype (het) rs2303067 (hom) | | 0.61 | 0.06 | 6.20 | .23915 |
| IC (A, B, E) | Interaction score* | 1 vs 0 | 1.08 | 0.72 | 1.61 | .71881 |
| | | 2 vs 0 | 1.18 | 0.80 | 1.76 | .39974 |
| | | 3 vs 0 | 1.67 | 1.08 | 2.57 | .01982 |
| | | 4 vs 0 | 2.75 | 1.53 | 4.92 | .00067 |
| | | 5 vs 0 | 1.45 | 0.14 | 14.70 | .75107 |
| MNM (A, E) | **Multinomial logit trait** | **Covariates** | **OR** | **95% CI** | | ***P* value** |
| | Response categories: 0: control subjects 1: eczema and *SPINK5* (not maternal) 2: eczema and *SPINK5* (maternal) | 1: sex, F vs M | 1.69 | 1.12 | 2.56 | .01338 |
| | | 1: *FLG* combined genotype | 5.22 | 3.12 | 8.73 | $4.9 \times 10{-}7$ |
| | | 2: sex, F vs M | 1.32 | 0.87 | 2.00 | .18653 |
| | | 2: *FLG* combined genotype | 6.16 | 3.70 | 10.26 | $1.1 \times 10{-}7$ |
| | Effect of *FLG* combined genotype in response category 1 compared to category 2 | | | | | .80071 |
| **Irish** | | | | | | |
| LRM (C, F) | **Gene-gene interaction term** | | **OR** | **95% CI** | | ***P* value** |
| IC (C, F) | *FLG* combined genotype (het vs wt) | | 4.34 | 2.77 | 6.79 | $1.3 \times 10^{-10}$ |
| | Interaction score* | 1 vs 0 | 1.32 | 0.64 | 2.75 | .45470 |
| | | 2 vs 0 | 1.18 | 0.56 | 2.46 | .66276 |
| | | 3 vs 0 | 1.96 | 0.92 | 4.20 | .08175 |
| | | 4 vs 0 | 5.17 | 1.88 | 14.24 | .00148 |
| | | 5 vs 0 | 0.01 | ~0 | $2.6 \times 10^{10}$ | .74599 |
| **ALSPAC** | | | | | | |
| LRM (D, G) | **Gene-gene interaction term** | | **OR** | **95% CI** | | ***P* value** |
| | *FLG* combined genotype (het vs wt) | | 2.26 | 1.63 | 3.14 | $1.2 \times 10^{-6}$ |
| | rs2303067 (het vs wt) | | 1.02 | 0.87 | 1.19 | .83390 |
| | rs2303067 (hom vs wt) | | 1.18 | 0.98 | 1.41 | .07389 |
| | *FLG* combined genotype (het) rs2303067 (het) | | 0.90 | 0.59 | 1.38 | .63859 |
| | *FLG* combined genotype (het) rs2303067 (hom) | | 1.04 | 0.62 | 1.74 | .87865 |
| IC (D, G) | Interaction score* | 1 vs 0 | 1.06 | 0.92 | 1.23 | .39907 |
| | | 2 vs 0 | 1.29 | 1.09 | 1.51 | .00245 |
| | | 3 vs 0 | 2.72 | 1.84 | 4.02 | $5.4 \times 10^{-7}$ |
| **Pooled†** | | | | | | |
| LRM (A, B, C, D, E, F, G) | **Gene-gene interaction term** | | **OR** | **95% CI** | | ***P* value** |
| | *FLG* combined genotype (het vs wt) | | 3.43 | 2.71 | 4.35 | $1.7 \times 10^{-24}$ |
| | rs2303067 (het vs wt) | | 1.03 | 0.92 | 1.16 | .58801 |
| | rs2303067 (hom vs wt) | | 1.15 | 1.00 | 1.32 | .04914 |
| | *FLG* combined genotype (het) rs2303067 (het) | | 0.81 | 0.60 | 1.09 | .16864 |
| | *FLG* combined genotype (het) rs2303067 (hom) | | 0.96 | 0.67 | 1.37 | .82037 |
| IC (A, B, C, D, E, F, G) | Interaction score* | 1 vs 0 | 1.10 | 0.99 | 1.23 | .08116 |
| | | 2 vs 0 | 1.37 | 1.21 | 1.55 | $6.5 \times 10^{-7}$ |
| | | 3 vs 0 | 3.66 | 2.83 | 4.75 | $1.2 \times 10^{-22}$ |

*LRM*, Logistic regression model with product-interaction terms; *het*, heterozygous; *wt*, wild-type; *hom*, homozygous; *IC*, logistic regression model with interaction score; *MNM*, multinomial logistic model considering maternal inheritance; *F*, female; *M*, male.

All models are adjusted for age and sex. Every block refers to a single model. Populations as designated in Table E1 are indicated in parentheses.

*Score: number of copies of the "risk" alleles in the *FLG* combined genotype, *SPINK5* rs2303067, and *KLK7*. For ALSPAC and pooled analysis, the score is reduced to the number of copies of *FLG* and *SPINK5* alleles.

†For the pooled analyses in all models, we estimated a population effect as a confounder with *P* values of less than $10^{-4}$.

**TABLE IV.** Results for random forests with 2 randomly preselected variables in each split

| | German (A, B, E) | | Irish (C, F) | | ALSPAC (D, G) | | Pooled (A, B, C, D, E, F, G) | |
|---|---|---|---|---|---|---|---|---|
| RF | Mean − 2 SE | Mean − 2 SE | Mean − 2 SE | Mean + 2 SE | Mean − 2 SE | Mean + 2 SE | Mean −2 SE | Mean + 2 SE |
| Age | 0.12179 | 0.12241 | 0.19797 | 0.19891 | NA | NA | NA | NA |
| Sex | −0.00037 | −0.00032 | 0.01075 | 0.01106 | −0.00004 | −0.00003 | −0.00022 | −0.00021 |
| *FLG* combined genotype | 0.00465 | 0.00473 | 0.05204 | 0.05253 | 0.00001 | 0.00002 | 0.00468 | 0.00474 |
| *SPINK5* | −0.00021 | −0.00016 | 0.00040 | 0.00059 | −0.00003 | −0.00002 | −0.00024 | −0.00022 |
| *KLK7* | −0.00015 | −0.00011 | −0.00112 | −0.00097 | NA | NA | NA | NA |
| Population | NA | NA | NA | NA | NA | NA | −0.00008 | −0.00006 |

*RF*, Random forest; *NA*, not available.

The average permutation importance ± 2 SEM over 100 iterations is displayed for each variable. Results for random forests with 5 randomly preselected variables (ie, for bagging) were almost identical. High positive values indicate a high variable importance. Small positive or negative values indicate that a variable is irrelevant for predicting the response. Populations as designated in Table I are indicated in parentheses.

predicting the response. The most commonly used variable importance measure is the permutation importance.[49]

High positive values of the importance measure obtained indicate a high variable importance. Small positive or negative values indicate that a variable is irrelevant for predicting the response.

Additionally, we pooled all 3 study cohorts to increase the power of detecting any single SNP and interaction effect. Because British and German populations might be slightly different with regard to ethnicity,[17] we accounted for a population effect in every analysis approach by introducing a binary independent variable, which codes as follows: 1, British/Irish origin; 0, German origin. Thus we corrected the estimated genetic effect for potential population differences.

Power calculations[50] for the pooled single SNP analyses were performed with nQuery7.0, assuming a dominant model.

All statistical analyses were carried out with R 2.6.0 (R Foundation for Statistical Computing, Vienna, Austria),[51] unless otherwise stated.

## RESULTS

### Single-gene analyses

First we examined the effect of the individual polymorphisms in predisposition to eczema in our samples. Allele frequencies are presented in Table E3. The *SPINK5* and *KLK7* polymorphisms showed comparable allele frequencies across all study populations. *FLG* polymorphism results have been published for the German family study previously,[9,17] although for this study, further families and patients were tested to increase statistical power, in particular when looking for interaction between alleles. In the German family cohort *FLG* polymorphisms greatly increased the risk for eczema (OR, 2.75; 95% CI, 1.93-3.98; $P = 1.8 \times 10^{-8}$), whereas the *SPINK5* polymorphism rs2303067 showed only a slight overtransmission to eczema-affected offspring (OR, 1.25; 95% CI, 1.04-1.50; $P = .018$). No associations were seen for the *KLK7* 3′UTR insertion (Table I). In this family cohort the power to detect a proportion of 1.2 between transmission of the risk allele versus nontransmission (which corresponds to a difference in the proportions of 5% given the observed number of discordant pairs) was 35% for the *SPINK5* polymorphism and 44% for the *KLK7* insertion. The power to detect a difference in the proportion of 9% for *SPINK5* and of 8% for *KLK7* was 80%.

Because parent-of-origin effects have been reported for *SPINK5* variants,[36] we also tested for differences between maternal and paternal allele sharing, and we confirmed a stronger association for the maternally inherited rs23030067 A allele and a relative risk of transmission of maternal alleles compared with paternal alleles of 2.18 (95% CI, 1.38-3.43; $P = .0008$). The observed power to detect parent-of-origin effects, as described in Weinberg,[45] was greater than 90%. In the pooled German case-control cohort, the presence of at least 1 *FLG* variant greatly increased the risk for eczema (OR, 4.56; 95% CI, 3.44-6.05; $P = 5.9 \times 10^{-26}$). In contrast, neither the *SPINK5* nor the *KLK7* variant was associated with eczema (Table II).

In the Irish case-control cohort comparable results could be seen: presence of an *FLG*-null allele increased the risk for eczema about 5.88-fold (95% CI, 3.85-8.99; $P = 2.8 \times 10^{-16}$). No association between *SPINK5* or *KLK7* and eczema was detected (Table II).

The analyses of the ALSPAC cohort also showed a strong *FLG* effect (OR, 2.23; 95% CI, 1.87-2.67; $P = 1.2 \times 10^{-18}$). Consistent with our other cohorts, no association between the Lys420Ser *SPINK5* polymorphism and eczema was found (Table II).

For the pooled study in all models, a population effect was estimated as a confounder with $P$ values of less than $10^{-4}$. An exceptionally strong *FLG* effect on eczema was seen for the pooled analyses (OR, 3.36; 95% CI, 2.97-3.79; $P = 1.3 \times 10^{-84}$), with a power of greater than 99% for the observed proportion of *FLG* variants of 0.21 in the cases. No association was observed for the *SPINK5* variant. The power to detect an increased risk in OR of 1.2 for carriers of the *SPINK5* rare allele compared with noncarriers with an observed proportion of 0.73 in the cases was 95%, assuming a dominant genetic model. For *KLK7*, we had 78% power to detect an OR of 1.2 in the pooled analyses, with an observed minor allele frequency in the cases of 0.56.

### Gene-gene interaction analyses

Gene-gene interactions were examined in all 3 cohorts separately and together in a pooled analysis.

The logistic regression model with product-interaction terms approach in the German case-control cohort was the best-fitting model on the basis of minimization of AIC with main effects of *FLG* and rs2303067 *(SPINK5)* and their product-interaction terms. In this model only *FLG* showed a significant genetic effect. In the Irish case-control cohort only *FLG*, along with the covariates age and sex, remained in the model according to the AIC (Table III).

In the interaction score approach, in addition to 1 mutant *FLG* allele, the A allele of rs2303067 *(SPINK5)* and the *KLK7* insertion were defined as "risk" variants. By deriving a score from the numbers of variant copies in the German case-control cohort, we observed a tendency for an increasing risk with the number of risk alleles an individual carried. The decrease in OR in the last category can probably be attributed to the low number of

observations, as reflected by the wide CI. Interestingly, the effect size increases with the number of risk alleles, probably because of the increased chance of *FLG* risk alleles in these cells.

For the Irish/British case-control cohort, similar results were observed, but the OR in the 4-variant group was more than twice as high as in the German case-control analyses. Using the multinomial regression model approach, we tried to account for the maternal parent-of-origin effect reported by Walley et al.[36] In our family collection we observed a tendency for an increasing risk of development of eczema caused by *FLG* mutations (regardless of inheritance of these *FLG* mutations) if the *SPINK5* SNP rs2303067 was inherited from the mother. There was an increased risk of eczema compared with paternal inheritance. However, the null hypothesis for equal *FLG* risk in both response categories could not be rejected.

In the random forest approach (Table IV), in the German sample only age and *FLG* status showed positive variable importance values. The average out-of-bag prediction accuracy was between 87.7% for the random forests and 88.3% for bagging. However, this is due to the fact that the average specificity was close to 100%, whereas the sensitivity was around 43% in the sample with approximately 21% cases. In the Irish case-control cohort only age and, to some extent, *FLG* and sex were suitable for predicting eczema. The average out-of-bag prediction accuracy was between 83.2% for the random forests and 84.4% for bagging.

The pooled analyses revealed results consistent with the previously estimated effects. In all models we estimated a population effect as a confounder with $P$ values of less than $10^{-4}$.

## DISCUSSION

This large-scale study examined variants in 3 candidate genes that have previously been reported to be associated with eczema. All genes encode proteins that are involved in the highly organized process of epidermal differentiation and are important for the maintenance of skin barrier function. In addition, because of their biologic interactions, we hypothesized that there might also be gene-gene-interactions.

Filaggrin is a key protein for the development of the cornified envelope and the process of cornification.[52] Two common null mutations in the *FLG* gene (R501X and 2282del4) have been firmly established as strong risk factors for eczema.[22,23] *KLK7* encodes the protease SSCE, which has been suggested to be involved in the complex proteolytic processing of filaggrin. An insertion in the 3′UTR of the *KLK7* gene possibly influencing SSCE activity has been reported to be associated with eczema in a UK case-control study, but this association has not been replicated thus far.[14,35] *SPINK5* is the gene defective in Netherton syndrome and encodes the serine proteinase inhibitor LEKTI, which has been implicated in the regulation of SSCE activity. An association of a *SPINK5* SNP with eczema has previously been reported[36] but was not confirmed in a recent study.[53]

Using a large cohort of German families and an Irish/English case-control series, as well as a pooled and enlarged German case-control collection and the longitudinal ALSPAC cohort, we first examined the individual SNPs. Results from these analyses suggest that of the tested polymorphisms, only the *FLG* mutations represent important and replicable genetic determinants for eczema, whereas the *SPINK5* variant Glu420Lys appears to have a weaker effect and only when maternally inherited, and the *KLK7* insertion does not exert an effect. However, our data do not preclude that there are other variants in *SPINK5*, *KLK7*, or both of importance for eczema.

In a second step we aimed at elucidating potential gene-gene interaction. Using several statistical approaches, we found no evidence for an interaction between variants in these 3 genes. The fact that the random forests permutation importance is essentially zero for the *SPINK5* and *KLK7* variants confirms that these variants are not relevant for predicting eczema, neither individually nor in interactions with each other, because interactions would be captured by the random forest variable importance.[54] Because there was a considerable difference in age between patients and control subjects, we considered age as a covariate in all analyses and modeled it nonparametrically. However, this leads to little variability in the response data explained by *FLG* mutations in the random forest approach.

It is widely hypothesized that complex human diseases, such as eczema, result from an unknown number of genetic factors, each of which influences susceptibility through interactions with other genes and with environmental factors.[55,56] With whole-genome association studies with hundreds of thousands of measured genetic variations emerging, analyses of the complex molecular interactions on the DNA level are of utmost importance, and it will be necessary to develop innovative statistical methods. For eczema, this is the first study that directly addresses this issue by exploring the effect of potential interactions among genes encoding proteins in the filaggrin expression and processing pathways. Using diverse and complementary statistical approaches in this large sample, we did not find evidence for epistatic effects between *FLG* and *KLK7* variants that significantly predict eczema risk. Thus although our data underline the exceptional importance of *FLG* deficiency for eczema risk, it does not support the hypothesis that its effect is dependent on or modified by *KLK7*. The results of the pooled analyses and the family analyses give a hint that *SPINK5* might be a potential player (when maternally inherited) within the filaggrin cascade, but this association requires further exploration. However, it cannot be excluded that acquired alterations in filaggrin processing or variations in other genes in the same pathway, such as *KLK5*, might contribute to eczema susceptibility. Functional studies are needed to explore the individual roles of products of genes within the filaggrin pathway and their biologic interactions, and future large-scale studies with powerful statistical methods will aid in elucidating the relationship between combinations of polymorphisms for eczema susceptibility.

We thank all of the patients and families who took part in this study, the professionals who helped in recruiting them, and the KORA and ALSPAC teams, which include interviewers, computer and laboratory technicians, clerical workers, research scientists, volunteers, managers, receptionists, and nurses.

**Key messages**

- **The *SPINK5* Lys420Ser polymorphism confers a risk of eczema when maternally inherited but is not a major genetic contributor to eczema risk.**
- **A previously reported association of a *KLK7* insertion and eczema could not be confirmed.**
- **There is no evidence for epistatic effects between *KLK7* or *SPINK5* variants and *FLG* mutations.**

## REFERENCES

1. van Beijsterveldt CE, Boomsma DI. Genetics of parentally reported asthma, eczema and rhinitis in 5-yr-old twins. Eur Respir J 2007;29:516-21.
2. Morar N, Willis-Owen SA, Moffatt MF, Cookson WO. The genetics of atopic dermatitis. J Allergy Clin Immunol 2006;118:24-34; quiz 5-6.
3. Glazier AM, Nadeau JH, Aitman TJ. Finding genes that underlie complex traits. Science 2002;298:2345-9.
4. Smith FJ, Irvine AD, Terron-Kwiatkowski A, Sandilands A, Campbell LE, Zhao Y, et al. Loss-of-function mutations in the gene encoding filaggrin cause ichthyosis vulgaris. Nat Genet 2006;38:337-42.
5. Palmer CN, Irvine AD, Terron-Kwiatkowski A, Zhao Y, Liao H, Lee SP, et al. Common loss-of-function variants of the epidermal barrier protein filaggrin are a major predisposing factor for atopic dermatitis. Nat Genet 2006;38:441-6.
6. Sandilands A, O'Regan GM, Liao H, Zhao Y, Terron-Kwiatkowski A, Watson RM, et al. Prevalent and rare mutations in the gene encoding filaggrin cause ichthyosis vulgaris and predispose individuals to atopic dermatitis. J Invest Dermatol 2006;126:1770-5.
7. Ruether A, Stoll M, Schwarz T, Schreiber S, Folster-Holst R. Filaggrin loss-of-function variant contributes to atopic dermatitis risk in the population of Northern Germany. Br J Dermatol 2006;155:1093-4.
8. Marenholz I, Nickel R, Ruschendorf F, Schulz F, Esparza-Gordillo J, Kerscher T, et al. Filaggrin loss-of-function mutations predispose to phenotypes involved in the atopic march. J Allergy Clin Immunol 2006;118:866-71.
9. Weidinger S, Illig T, Baurecht H, Irvine AD, Rodriguez E, Diaz-Lacava A, et al. Loss-of-function variations within the filaggrin gene predispose for atopic dermatitis with allergic sensitizations. J Allergy Clin Immunol 2006;118:214-9.
10. Stemmler S, Parwez Q, Petrasch-Parwez E, Epplen JT, Hoffjan S. Two common loss-of-function mutations within the filaggrin gene predispose for early onset of atopic dermatitis. J Invest Dermatol 2007;127:722-4.
11. Barker JN, Palmer CN, Zhao Y, Liao H, Hull PR, Lee SP, et al. Null mutations in the filaggrin gene (FLG) determine major susceptibility to early-onset atopic dermatitis that persists into adulthood. J Invest Dermatol 2007;127:564-7.
12. Sandilands A, Terron-Kwiatkowski A, Hull PR, O'Regan GM, Clayton TH, Watson RM, et al. Comprehensive analysis of the gene encoding filaggrin uncovers prevalent and rare mutations in ichthyosis vulgaris and atopic eczema. Nat Genet 2007;39:650-4.
13. Hubiche T, Ged C, Benard A, Leaute-Labreze C, McElreavey K, de Verneuil H, et al. Analysis of SPINK 5, KLK 7 and FLG genotypes in a French atopic dermatitis cohort. Acta Derm Venereol 2007;87:499-505.
14. Morar N, Cookson WO, Harper JI, Moffatt MF. Filaggrin mutations in children with severe atopic dermatitis. J Invest Dermatol 2007;127:1667-72.
15. Nomura T, Sandilands A, Akiyama M, Liao H, Evans AT, Sakai K, et al. Unique mutations in the filaggrin gene in Japanese patients with ichthyosis vulgaris and atopic dermatitis. J Allergy Clin Immunol 2007;119:434-40.
16. Rogers AJ, Celedon JC, Lasky-Su JA, Weiss ST, Raby BA. Filaggrin mutations confer susceptibility to atopic dermatitis but not to asthma. J Allergy Clin Immunol 2007;120:1332-7.
17. Weidinger S, Rodriguez E, Stahl C, Wagenpfeil S, Klopp N, Illig T, et al. Filaggrin mutations strongly predispose to early-onset and extrinsic atopic dermatitis. J Invest Dermatol 2007;127:724-6.
18. Nomura T, Akiyama M, Sandilands A, Nemoto-Hasebe I, Sakai K, Nagasaki A, et al. Specific filaggrin mutations cause ichthyosis vulgaris and are significantly associated with atopic dermatitis in Japan. J Invest Dermatol 2008; 128:1436-41.
19. Ekelund E, Lieden A, Link J, Lee SP, D'Amato M, Palmer CN, et al. Loss-of-function variants of the filaggrin gene are associated with atopic eczema and associated phenotypes in Swedish families. Acta Derm Venereol 2008;88:15-9.
20. Brown SJ, Relton CL, Liao H, Zhao Y, Sandilands A, Wilson IJ, et al. Filaggrin null mutations and childhood atopic eczema: a population-based case-control study. J Allergy Clin Immunol 2008;121:940-6.
21. Baurecht H, Irvine AD, Novak N, Illig T, Buhler B, Ring J, et al. Towards a major risk factor for atopic eczema: meta-analysis of filaggrin mutation data. J Allergy Clin Immunol 2007;120:1406-12.
22. Rodriguez E, Illig T, Weidinger S. Filaggrin loss-of-function mutations and association with allergic diseases. Pharmacogenomics 2008;9:399-413.
23. Weidinger S, O'Sullivan M, Illig T, Baurecht H, Depner M, Rodriguez E, et al. Filaggrin mutations, atopic eczema, hay fever, and asthma in children. J Allergy Clin Immunol 2008;121:1203-9.
24. Henderson J, Northstone K, Lee SP, Liao H, Zhao Y, Pembrey M, et al. The burden of disease associated with filaggrin mutations: a population-based, longitudinal birth cohort study. J Allergy Clin Immunol 2008;121:872-7.
25. Hudson TJ. Skin barrier function and allergic risk. Nat Genet 2006;38:399-400.
26. McLean WH, Hull PR. Breach delivery: increased solute uptake points to a defective skin barrier in atopic dermatitis. J Invest Dermatol 2007;127:8-10.
27. Sicherer SH, Leung DY. Advances in allergic skin disease, anaphylaxis, and hypersensitivity reactions to foods, drugs, and insects. J Allergy Clin Immunol 2007;119: 1462-9.
28. Resing KA, Walsh KA, Dale BA. Identification of two intermediates during processing of profilaggrin to filaggrin in neonatal mouse epidermis. J Cell Biol 1984;99:1372-8.
29. Resing KA, Walsh KA, Haugen-Scofield J, Dale BA. Identification of proteolytic cleavage sites in the conversion of profilaggrin to filaggrin in mammalian epidermis. J Biol Chem 1989;264:1837-45.
30. Resing KA, Thulin C, Whiting K, al-Alawi N, Mostad S. Characterization of profilaggrin endoproteinase 1. A regulated cytoplasmic endoproteinase of epidermis. J Biol Chem 1995;270:28193-8.
31. Descargues P, Deraison C, Bonnart C, Kreft M, Kishibe M, Ishida-Yamamoto A, et al. Spink5-deficient mice mimic Netherton syndrome through degradation of desmoglein 1 by epidermal protease hyperactivity. Nat Genet 2005;37:56-65.
32. Komatsu N, Takata M, Otsuki N, Ohka R, Amano O, Takehara K, et al. Elevated stratum corneum hydrolytic activity in Netherton syndrome suggests an inhibitory regulation of desquamation by SPINK5-derived peptides. J Invest Dermatol 2002;118:436-43.
33. Bitoun E, Micheloni A, Lamant L, Bonnart C, Tartaglia-Polcini A, Cobbold C, et al. LEKTI proteolytic processing in human primary keratinocytes, tissue distribution and defective expression in Netherton syndrome. Hum Mol Genet 2003;12:2417-30.
34. Vasilopoulos Y, Cork MJ, Murphy R, Williams HC, Robinson DA, Duff GW, et al. Genetic association between an AACC insertion in the 3'UTR of the stratum corneum chymotryptic enzyme gene and atopic dermatitis. J Invest Dermatol 2004;123:62-6.
35. Walley AJ, Chavanas S, Moffatt MF, Esnouf RM, Ubhi B, Lawrence R, et al. Gene polymorphism in Netherton and common atopic disease. Nat Genet 2001;29:175-8.
36. Kato A, Fukai K, Oiso N, Hosomi N, Murakami T, Ishii M. Association of SPINK5 gene polymorphisms with atopic dermatitis in the Japanese population. Br J Dermatol 2003;148:665-9.
37. Nishio Y, Noguchi E, Shibasaki M, Kamioka M, Ichikawa E, Ichikawa K, et al. Association between polymorphisms in the SPINK5 gene and atopic dermatitis in the Japanese. Genes Immun 2003;4:515-7.
38. Rathmann W, Haastert B, Icks A, Lowel H, Meisinger C, Holle R, et al. High prevalence of undiagnosed diabetes mellitus in Southern Germany: target populations for efficient screening. The KORA survey 2000. Diabetologia 2003;46:182-9.
39. Burney PG, Luczynska C, Chinn S, Jarvis D. The European Community Respiratory Health Survey. Eur Respir J 1994;7:954-60.
40. Hanifin JM, Rajka G. Diagnostic features of atopic eczema. Acta Derm Venereol 1980;92:44-7.
41. Williams HC, Burney PG, Hay RJ, Archer CB, Shipley MJ, Hunter JJ, et al. The U.K. Working Party's Diagnostic Criteria for Atopic Dermatitis. I. Derivation of a minimum set of discriminators for atopic dermatitis. Br J Dermatol 1994;131:383-96.
42. Golding J, Pembrey M, Jones R. ALSPAC—the Avon Longitudinal Study of Parents and Children. I. Study methodology. Paediatr Perinat Epidemiol 2001;15:74-87.
43. Jones RW, Ring S, Tyfield L, Hamvas R, Simmons H, Pembrey M, et al. A new human genetic resource: a DNA bank established as part of the Avon longitudinal study of pregnancy and childhood (ALSPAC). Eur J Hum Genet 2000;8:653-60.
44. Weinberg CR. Methods for detection of parent-of-origin effects in genetic studies of case-parents triads. Am J Hum Genet 1999;65:229-35.
45. Foster MR. Key concepts in model selection: Performance and generalizability. J Math Psychol 2000;44:205-31.
46. Weedon MN, McCarthy MI, Hitman G, Walker M, Groves CJ, Zeggini E, et al. Combining information from common type 2 diabetes risk polymorphisms improves disease prediction. PLoS Med 2006;3:e374.
47. Brezger A, Kneib T, Lang S. BayesX—software for Bayesian inference in structured additive regression models. Munich: Department for Statistics; 2007. Available at: http://www.stat.uni-muenchen.de/~bayesx/bayesx.html. Accessed July 25, 2008.
48. McCullagh P, Nelder JA. Generalized linear models. New York (NY): Chapman and Hall; 1989.
49. Strobl C, Boulesteix AL, Zeileis A, Hothorn T. Bias in random forest variable importance measures: illustrations, sources and a solution. BMC Bioinformatics 2007;8:25.
50. Fleiss JL, Tytun A, Ury HK. A simple approximation for calculating sample sizes for comparing independent proportions. Biometrics 1980;36:343-6.
51. Team RDC. R: a language and environment for statistical computing. Vienna: R Foundation for Statistical Computing; 2007.
52. Candi E, Schmidt R, Melino G. The cornified envelope: a model of cell death in the skin. Nat Rev Mol Cell Biol 2005;6:328-40.
53. Folster-Holst R, Stoll M, Koch WA, Hampe J, Christophers E, Schreiber S. Lack of association of SPINK5 polymorphisms with nonsyndromic atopic dermatitis in the population of Northern Germany. Br J Dermatol 2005;152:1365-7.
54. Lunetta KL, Hayward LB, Segal J, Van Eerdewegh P. Screening large-scale association study data: exploiting interactions using random forests. BMC Genet 2004;5:32.
55. Moore JH. The ubiquitous nature of epistasis in determining susceptibility to common human diseases. Hum Hered 2003;56:73-82.
56. Moore JH. A global view of epistasis. Nat Genet 2005;37:13-4.

## METHODS

### Statistical methods

Descriptive statistics for quantitative and qualitative values are presented as means ± SD and relative frequencies or absolute numbers, respectively. Deviation from Hardy-Weinberg equilibrium was tested in parents for the family analyses and in control subjects for the case-control analyses. In the family setting we analyzed association of single SNPs with eczema using the classical transmission disequilibrium test. Parent-of-origin effects were investigated with the method proposed by Weinberg.[E1]

Case-control analyses for single SNPs were performed by using logistic regression models adjusted for age and sex. To not constrain the analyses to a specific genetic model, we modeled the 3 categorical genotypes by 2 dummy variables. Thus we estimated separate effects for heterozygotes and homozygotes compared with the wild-type. Because the English and Irish case series are ethnically related, were shown to have highly similar genotypes for *FLG*,[E2,E3] and here had almost identical minor allele frequencies for *KLK7* and *SPINK5* polymorphisms, we analyzed these cases together.

Gene-gene interaction analyses were performed after excluding individuals with 2 mutant *FLG* alleles because these individuals do not express the filaggrin protein, and therefore a biologic interaction with posttranslational modifications by SSCE and LEKTI is not plausible.

Four different approaches were carried out to estimate interaction effects between the single polymorphisms in *FLG*, *KLK7*, and *SPINK5*. In any of these approaches, we adjusted the models for the common covariates age and sex. First, interaction was evaluated by using the logistic regression model with product-interaction terms. We started with a sparse model of the known covariates age, sex, and *FLG*. Incrementally, we extended the model with additional SNPs in *SPINK5* and *KLK7* and SNP-SNP product-interaction terms. The AIC was used to select the appropriate model.[E4]

Second, we defined an interaction score, which counts the number of copies of the potentially disease-associated alleles,[E5] which we used as a covariate in the logistic model.

Third, we modeled maternal effects observed for *SPINK5* using affected offspring from families only. We constructed a 3-category trait: unaffected control subjects, affected offspring with no mutant allele inherited from the mother, and affected offspring with a mutant allele inherited from the mother. We then estimated ORs for both affected status compared with control subjects in a multinomial regression model.

For any of these regression approaches, the quantitative covariate age was modeled nonparametrically in a general additive model framework because of the functional structure of Anscombe residuals after applying analyses in the general linear model framework. For the multinomial model, we used a restricted maximum likelihood approach[E6] implemented in BayesX 1.50.[E7]

Finally, for further exploration, variable importance measures were computed by means of the random forest method. Random forests, and the related method of bagging, are an ensemble method in which a set of classification or regression trees is aggregated for prediction.[E8,E9] Random forests provide variable importance measures that can be used to detect variables relevant for predicting the response. The most commonly used variable importance measure is the permutation importance.[E10]

For variable selection purposes, the advantage of the random forest permutation accuracy importance measure compared with univariate screening methods is that it covers the effect of each predictor variable individually, as well as in multivariate interactions with other predictor variables. For example, Lunetta et al[E11] demonstrated that genetic markers relevant in interactions with other markers or environmental variables can be detected more efficiently by means of random forests than by means of univariate screening methods, such as the Fisher exact test. Here the random forest implementation cforest from the package party[E12,E13] in the R system for statistical computing[E14] is used because it guarantees unbiased variable selection for predictor variables of different scales of measurement.[E10] Predictor variables considered here were age, sex, and respective SNP variables.

One hundred random forests with 500 trees each were fitted with the configuration guaranteeing unbiased variable selection suggested by Strobl et al[E10] to assess the stability of the results. The random forests were built with either 2 randomly preselected variables in each split (argument mtry = 2) or 3 randomly preselected variables in each split (mtry = 3) for comparison. The latter approach is equivalent to bagging, which is contained in random forests as the special case in which the number of randomly preselected variables is equal to the number of available variables. High positive values of the importance measure obtained indicate a high variable importance. Small positive or negative values indicate that a variable is irrelevant for predicting the response.

Additionally, we pooled all 3 study cohorts to increase the power of detecting any single SNP and interaction effect. Because British and German populations might be slightly different with regard to ethnicity,[E15] we accounted for a population effect in every analysis approach by introducing a binary variable, which codes as follows: 1, British/Irish origin; 0, German origin. Thus we corrected the estimated genetic effect for potential population differences. Power calculations[E16] for the pooled single SNP analyses were performed with nQuery7.0, assuming a dominant model.

All statistical analyses were carried out with R 2.6.0,[E14] unless otherwise stated.

#### REFERENCES

E1. Weinberg CR. Methods for detection of parent-of-origin effects in genetic studies of case-parents triads. Am J Hum Genet 1999;65:229-35.
E2. Sandilands A, O'Regan GM, Liao H, Zhao Y, Terron-Kwiatkowski A, Watson RM, et al. Prevalent and rare mutations in the gene encoding filaggrin cause ichthyosis vulgaris and predispose individuals to atopic dermatitis. J Invest Dermatol 2006;126:1770-5.
E3. Barker JN, Palmer CN, Zhao Y, Liao H, Hull PR, Lee SP, et al. Null mutations in the filaggrin gene (FLG) determine major susceptibility to early-onset atopic dermatitis that persists into adulthood. J Invest Dermatol 2007;127:564-7.
E4. Foster MR. Key concepts in model selection: performance and generalizability. J Math Psychol 2000;44:205-31.
E5. Weedon MN, McCarthy MI, Hitman G, Walker M, Groves CJ, Zeggini E, et al. Combining information from common type 2 diabetes risk polymorphisms improves disease prediction. PLoS Med 2006;3:e374.
E6. Kneib T, Baumgartner B, Steiner WJ. Semiparametric multinomial logit models for analysing consumer choice behaviour. AStA Adv Stat Anal 2007;91:225-44.
E7. Brezger A, Kneib T, Lang S. BayesX—software for Bayesian inference in structured additive regression models. Munich: Department for Statistics; 2007. Available at: http://www.stat.uni-muenchen.de/~bayesx/bayesx.html. Accessed July 25, 2008.
E8. Breiman L. Arcing classifiers. Ann Stat 1998;26:801-49.
E9. Breiman L. Random forests. Machine Learn 2001;45:5-32.
E10. Strobl C, Boulesteix AL, Zeileis A, Hothorn T. Bias in random forest variable importance measures: illustrations, sources and a solution. BMC Bioinformatics 2007;8:25.
E11. Lunetta KL, Hayward LB, Segal J, Van Eerdewegh P. Screening large-scale association study data: exploiting interactions using random forests. BMC Genet 2004;5:32.
E12. Hothorn T, Hornik K, Zeileis A. Unbiased recursive partitioning: a conditional inference framework. J Comp Graph Stat 2006;15:651-74.
E13. Hothorn T, Hornik K, Zeileis A. Party: a laboratory for recursive part(y)itioning package version 0.9-0. 2006.
E14. Team RDC. R: a language and environment for statistical computing. Vienna: R Foundation for Statistical Computing; 2007.
E15. Rogers AJ, Celedon JC, Lasky-Su JA, Weiss ST, Raby BA. Filaggrin mutations confer susceptibility to atopic dermatitis but not to asthma. J Allergy Clin Immunol 2007;120:1332-7.
E16. Fleiss JL, Tytun A, Ury HK. A simple approximation for calculating sample sizes for comparing independent proportions. Biometrics 1980;36:343-6.

**TABLE E1.** Descriptive characterization of case and control populations

| | n | Mean age (y [SD]) | Male sex | Mean IgE (SD) | Country | Notation |
|---|---|---|---|---|---|---|
| German offspring | 486 | 22.04 (10.64) | 198 (40.7%) | 990.4 (2472.7) | Germany | A |
| German cases | 287 | 35.55 (16.15) | 112 (39.0%) | 1442.2 (2380.6) | Germany | B |
| Irish/English cases | 418 | 19.42 (18.43) | 199 (51.4%) | 3008.0 (6170.0) | Ireland | C |
| ALSPAC cases* | 1583 | 3.5 | 849 (53.6%) | 286.0 (539.0) | England | D |
| Sum cases | 2774 | | | | | |
| KORA S4 | 3992 | 49.51 (13.90) | 1971 (49.9%) | 114.2 (1535.5) | Germany | E |
| Irish control subjects | 552 | 35.71 (12.27) | 170 (30.8%) | NA | Ireland | F |
| ALSPAC control subjects | 6063 | 3.5* | 3135 (51.7%) | 200.4 (462.7) | England | G |
| Sum control subjects | 10607 | | | | | |

*Eczema status was determined in all children at 42 months in the ALSPAC cohort.

**TABLE E2.** Genotyping details and minor allele frequencies in the KORA S4 population-based cohort

| SNP ID | MAF | DIR | PCR primer | Extension primer |
|---|---|---|---|---|
| R501X | 0.013 | fwd | ACGTTGGATGCTGGAGGAAGACAAGGATCG | ATGCCTGGAGCTGTCTC |
| | | rev | ACGTTGGATGATGGTGTCCTGACCCTCTTG | |
| 2282del4 | 0.025 | fwd | ACGTTGGATGTTGGTGGCTCTGCTGATGGT | GAAGACTCAGACACACAGT |
| | | rev | ACGTTGGATGGTGAGGGACATTCAGAAGAC | |
| rs2303067 | 0.479 | fwd | ACGTTGGATGCCATCCTTTTTTAGCCAAGC | GATTGTCTTTTGTTTCTTGATT |
| | | rev | ACGTTGGATGCCTCAAAGGAAGCTGTACTC | |
| AACC ins | 0.312 | fwd | ACGTTGGATGTGATTGGTTTATCAACAGG | TTTCCTCAAAGATATATTTAAACC |
| | | rev | ACGTTGGATGGACGCCGATGACCTATGAAG | |

*MAF*, Minor allele frequency; *DIR*, direction; *fwd*, forward; *rev*, reverse.

**TABLE E3.** Allele frequencies across study populations

| | GER | | IRL | | ALSPAC | |
|---|---|---|---|---|---|---|
| | Major | Minor | Major | Minor | Major | Minor |
| *SPINK5* | 0.5181 | 0.48186 | 0.5227 | 0.4773 | 0.5254 | 0.47463 |
| *KLK7* | 0.6891 | 0.31089 | 0.6612 | 0.3388 | NA | NA |
| *FLG* | 0.9480 | 0.05201 | 0.8639 | 0.1361 | 0.9558 | 0.04418 |

*NA*, Not applicable.

# Unique and Recurrent Mutations in the Filaggrin Gene in Singaporean Chinese Patients with Ichthyosis Vulgaris

Huijia Chen[1,2], Jean C.C. Ho[3], Aileen Sandilands[1], Yuin Chew Chan[3], Yoke Chin Giam[3], Alan T. Evans[4], E. Birgitte Lane[2] and W.H. Irwin McLean[1]

Filaggrin is an abundant protein of the outer epidermis that is essential for terminal differentiation of keratinocytes and formation of an effective barrier against water loss and pathogen/allergen/irritant invasion. Recent investigations in Europe and Japan have revealed null mutations in the filaggrin gene (*FLG*) as the underlying cause of ichthyosis vulgaris (IV), a common skin disorder characterised by dry skin, palmar hyperlinearity and keratosis pilaris. Following the development of a strategy for the comprehensive analysis of *FLG*, we have identified five unique mutations and one recurrent mutation in Singaporean Chinese IV patients. Mutation 441delA is located in the profilaggrin S100 domain, whereas two additional frameshift mutations, 1249insG and 7945delA, occur in the first partial filaggrin repeat ("repeat 0") and in filaggrin repeat 7, respectively. Both nonsense mutations Q2147X and E2422X are found in filaggrin repeat 6, whereas R4307X was found on one of the longer size variant alleles of *FLG*, within duplicated repeat 10.2. Mutation E2422X, previously found in a single Dutch patient, was found in one Singaporean IV patient and at a low frequency in Asian population controls. Our study confirms the presence of population-specific as well as recurrent *FLG* mutations in Singapore.

*Journal of Investigative Dermatology* (2008) **128,** 1669–1675; doi:10.1038/jid.2008.2; published online 31 January 2008

## INTRODUCTION

Ichthyosis vulgaris (IV; OMIM #146700) is one of the most common monogenic skin disorders, reported to affect 1 in 250 English school children (Wells, 1966). The disorder is characterized by dry, scaly skin, palmar hyperlinearity, and keratosis pilaris (Sybert *et al.*, 1985; Smith *et al.*, 2006), and can be aggravated by climate. Other allergic conditions, importantly including atopic dermatitis (AD), atopic asthma and rhinitis, are also commonly manifested by IV patients (Irvine and McLean, 2006; Irvine, 2007).

In 1985, Sybert *et al.* demonstrated that profilaggrin and filaggrin were reduced or absent in five IV patients. The filaggrin gene (*FLG*) is located on chromosome 1q21 within a cluster of genes that make up the EDC—the epidermal differentiation complex (Mischke *et al.*, 1996). Profilaggrin consists of three exons (Presland *et al.*, 1992)—exon 1 (15 bp) and exon 2 (159 bp) contain the 5′-untranslated region and start codon, respectively. Exon 3 is unusually large (12.7–14.7 kb) and codes for most of the polyprotein profilaggrin, which consists of 10–12 tandem repeats of the 37-kDa filaggrin peptide. Profilaggrin is stored within keratohyalin granules in the granular layer of the epidermis and proteolytically cleaved to filaggrin subunits when keratinocytes undergo terminal differentiation. Subsequently, filaggrin binds to keratin intermediate filaments within the keratinocytes, as the cells collapse and flatten into squames; transglutaminases also work to cross-link other cornified envelope precursor proteins such as involucrin to form an impermeable epidermal barrier. It has also been shown that filaggrin subunits are ultimately proteolyzed within the stratum corneum into hygroscopic amino acids and derivatives thereof, which help retain epidermal moisture (Rawlings and Harding, 2004).

Although filaggrin expression was known to be reduced or absent in IV patients for 20 years, the route to uncovering the direct genetic cause of IV was challenging because it was very difficult to conduct routine and effective sequencing of *FLG* due to its sheer size. The *FLG* gene is also highly repetitive and polymorphic, which prevents sequencing of the entire gene by conventional methods. In 2006, the precise molecular basis of IV was finally uncovered by the identification of null mutations in *FLG* (Smith *et al.*, 2006). Since IV was known to be associated with a high incidence of AD (Kuokkanen, 1969; Tay *et al.*, 2002), the link between

[1]*Epithelial Genetics Group, Human Genetics Unit, Division of Pathology and Neuroscience, University of Dundee, Ninewells Hospital and Medical School, Dundee, UK;* [2]*A*STAR Institute of Medical Biology, Singapore, Singapore;* [3]*National Skin Centre, Singapore, Singapore and* [4]*Department of Pathology, Tayside University Hospitals NHS Trust, Dundee, UK*

*Correspondence: Professor W.H. Irwin McLean, Epithelial Genetics Group, Human Genetics Unit, Division of Pathology and Neuroscience, University of Dundee, Ninewells Hospital and Medical School, Dundee DD1 9SY, UK. E-mail: w.h.i.mclean@dundee.ac.uk*

*Abbreviations: AD, atopic dermatitis;* FLG, *filaggrin gene; IV, ichthyosis vulgaris*

*Received 21 September 2007; revised 13 November 2007; accepted 19 November 2007; published online 31 January 2008*

filaggrin-null mutations and AD was quickly established (Palmer *et al.*, 2006). Reduced or absent filaggrin expression chronically exposes the skin to irritants that can lead to inflammation of skin, resulting in AD (Irvine, 2007). In some populations, up to 50% of children with moderate to severe AD carry *FLG* mutations and defects in *FLG* also predispose AD patients to asthma and related allergic phenotypes (Irvine and McLean, 2006).

This laboratory has recently reported methods for the comprehensive analysis of *FLG* (Sandilands *et al.*, 2007). Three different sizes of *FLG* alleles were identified containing 10, 11, or 12 filaggrin repeats. By overlapping PCR and unidirectional deletion strategies, *FLG* was fully sequenced, and a total of 15 null mutations in *FLG* in the European, Japanese and Chinese populations have been identified to date (Nomura *et al.*, 2007; Sandilands *et al.*, 2007). Interestingly, common mutations in the European population differ from those in the Oriental populations studied so far.

Singapore is an immigrant South-East Asian city-state with a population of 4.5 million consisting of Chinese (75.6%), Malays (13.6%), Indians (8.7%), and other ethnic groups (2.1%). The 1-year prevalence of IV in 2002 was reported to be 8% in Singaporean school children (Tay *et al.*, 2002). Similar to many developed European countries, AD cases are also on the rise in Singapore; a study conducted on 2,363 students in 2002 reported a prevalence of 20.8% (Tay *et al.*, 2002). In this study, we aimed to identify the *FLG* mutations that underlie IV in this Asian population to facilitate future studies of AD in this geographical region.

## RESULTS

### Case reports

All patients, adults of Chinese Singaporean ethinicity, were referred to outpatient dermatological clinics at the National Skin Centre, Singapore for the main complaint of dry skin, with varying accompanying symptoms of eczema. The diagnosis of IV was established on the clinical basis of finding extensive semi-adherent scaling, which was characteristic, a history of onset early in life; as well as a positive family history of a similar affliction. These criteria were chosen to increase the yield of a positive result on genetic testing. Palmar hyperlinearity was seen in all patients (Figure 1, Figure S1). Interestingly, keratosis pilaris was not seen at all. IV was graded in terms of severe or mild, based on the degree and extent (acral and truncal sites versus acral sites only respectively) of scaling (Figure 1). Four out of six patients with severe IV also had AD based on the widely recognized diagnostic criteria (Hannifin and Rajka, 1980). None of the patients with mild IV had active AD. All subjects were examined by a single dermatologist.

### Unique *FLG* mutations in Singaporean Chinese IV patients

We carried out full sequencing of the *FLG* in eight unrelated Singaporean Chinese IV patients and detected mutations in six individuals (Figures 3 and 4). A total of six mutations were detected (Table 2)—five mutations are previously unreported and one mutation was identified in a single Dutch IV patient from an earlier study (Sandilands *et al.*, 2007).

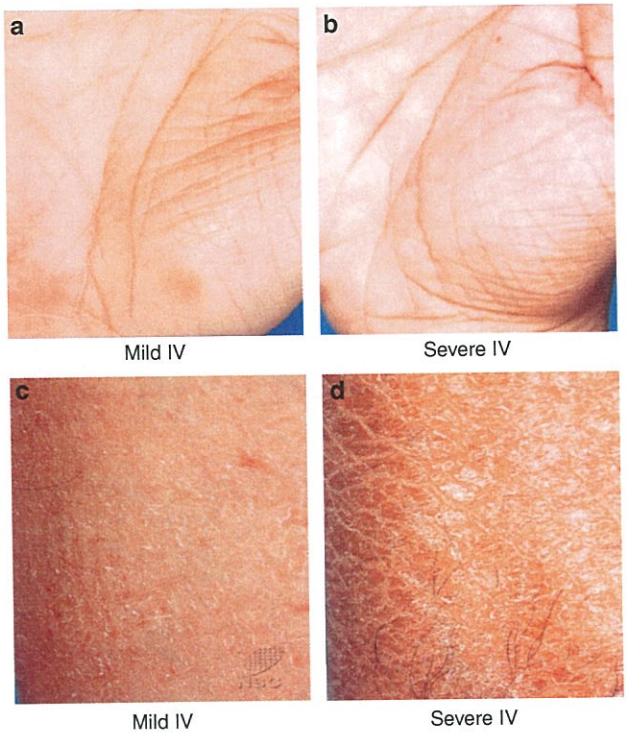



**Figure 1. Clinical features of patients with IV.** (**a**) Palmar hyperlinearity seen in patient 3 heterozygous for *FLG* mutation 7945delA. (**b**) Due to the semi-dominant inheritance pattern of IV, a more marked hyperlinearity can be seen on the palms of patient 8, who is a compound heterozygote for *FLG* mutations 441delA and 1249delG. (**c**) Fine scaling on limbs was also present in patient 3. (**d**) In contrast, prominent scaling on the trunk and limbs was seen in patients diagnosed with severe IV.

We identified a mutation within the S100 domain of profilaggrin, 441delA, which was detected in patient 8 and resulted in the introduction of a premature stop codon at position 193. In patients 5 and 8, mutation 1249insG was found in the first partial filaggrin repeat (termed "repeat 0") preceding the first complete filaggrin repeat, leading to a downstream premature stop codon at position 418. In repeat 6, two different nonsense mutations were detected in patients 4 and 6. Mutation E2422X and Q2417X mutations resulted from a single nucleotide base substitution of G>T or C>T at positions 7,249 and 7,264, respectively. Interestingly, mutation E2422X was previously reported in a Dutch IV patient (Sandilands *et al.*, 2007) and we also detected one heterozygous carrier in a cohort of 164 normal controls from China. Frameshift mutation 7945delA was detected in filaggrin repeat 7 in patients 3 and 5. Patient 2 harbored a nonsense mutation in repeat 10.2, the first mutation to be reported in one of the duplicated repeats found within the longer size variant alleles of *FLG* (Sandilands *et al.*, 2007). All mutations were confirmed by direct sequencing of PCR products. In addition, we further confirmed the presence of frameshift mutations (441delA, 1249insG, 7945delA) by sizing of fluorescently labeled PCR products and the three nonsense mutations (Q2417X, E2422X, R4307X) were also confirmed

by restriction digestion of PCR products. No *FLG* mutations were found in patients 1 and 7 (Table 2).

Screening of 100 Singaporean and 160 Chinese normal controls for the six mutations identified in Singaporean Chinese IV patients was also performed. With the exception of one case of E2422X found in a Chinese control, none of the other mutations were detected in the controls.

**Absent or reduced profilaggrin expression in IV patients**

Preliminary analysis of skin biopsies from all eight patients was conducted using hematoxylin and eosin staining. The appearance of keratohyalin granules was greatly reduced in the granular layer of all patients except patient 1 (Figure 2). To detect a change in filaggrin levels, we carried out immunohistochemical analysis of the skin biopsies with monoclonal antibody 15C10, which identifies an epitope conserved in all or most filaggrin repeats. Filaggrin staining in Patient 1 was comparable to that of normal skin (not shown). In contrast, filaggrin staining was greatly reduced in patients 2, 3, 4, and 6; barely detectable staining was observed in patients 5 and 8 (Figure 2), and staining was completely absent in patient 7 (not shown).

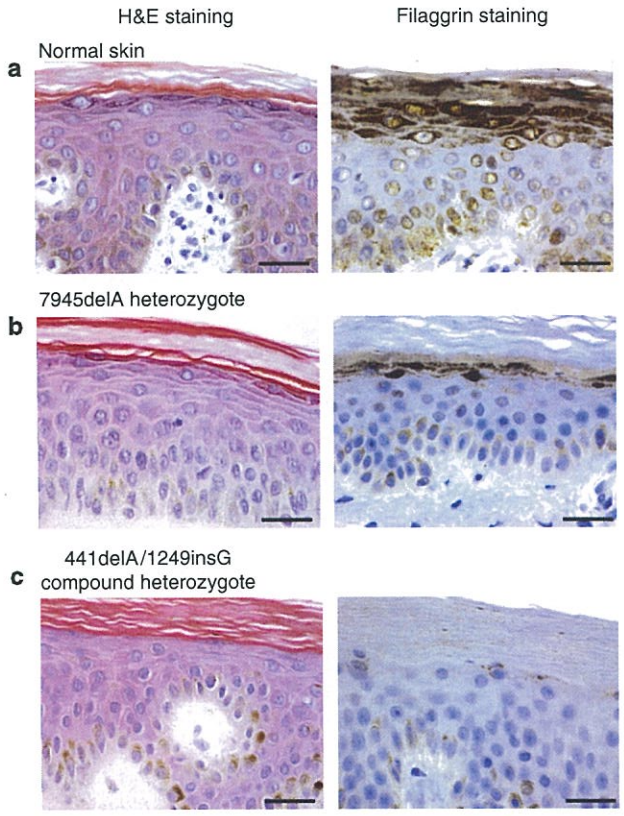



**Figure 2. Histological features of patients with IV.** All left panels show hematoxylin and eosin (H&E) staining and right panels show immunohistochemical staining with the 15C10 monoclonal antibody against filaggrin repeats in skin biopsies. (**a**) Keratohyalin granules are clearly visible in normal control skin after H&E staining. Immunohistochemical staining detects filaggrin repeats readily in 4–5 granular cell layers and throughout the stratum corneum. (**b**) The presence of a heterozygous 7945delA *FLG* mutation in patient 3 leads to reduced keratohyalin granules and filaggrin staining in the granular layer. The stratum corneum is also thickened. Patients 2, 4, and 5 show similar histochemical features. (**c**) Patient 8, who is a compound heterozygote for 441delA/1249insG *FLG* mutations, has barely detectable keratohyalin granules and filaggrin staining in the granular layer. The stratum corneum is severely thickened. Patient 6 shows similar histochemical features. Bars = 50 μm.

## DISCUSSION

Profilaggrin is an unusually large (12.7–14.7 kb) and repetitive gene that has proved to be very challenging to sequence for detection of pathogenic mutations. Recently, comprehensive re-sequencing primers have been developed to allow complete routine analysis of the entire *FLG* (Sandilands *et al.*, 2007).

In this report, we detected six mutations in Singaporean Chinese IV patients, of which five have not been seen in other populations previously studied (Gruber *et al.*, 2006; Marenholz *et al.*, 2006; Smith *et al.*, 2006; Barker *et al.*, 2007; Nomura *et al.*, 2007; Sandilands *et al.*, 2007; Stemmler *et al.*, 2007). Given the similar ancestry among the Singaporean Chinese, Chinese, and Japanese, we were surprised that the mutations previously detected in those populations were not found in our study cohort. The Chinese and Japanese populations share a common mutation 3321delA that is absent in the European population but interestingly, this was not found in our Singaporean Chinese patients despite being found in 3.6% of Chinese controls (Sandilands *et al.*, 2007). It may be that this mutation is less common in Southern Asian populations. With the known mutations now numbering more than 20, it is possible that *FLG* is constantly under evolutionary mutation pressure, so that every population will have a unique set of *FLG* mutations.

The five unique mutations reported here were not detected in any of the 260 normal control samples, therefore these mutations are probably rare. This could also explain why these mutations were not previously reported in the Japanese and Chinese patients studied previously (Nomura *et al.*, 2007; Sandilands *et al.*, 2007). Alternatively, these mutations may also have arisen quite recently amongst the Singaporean Chinese. Interestingly, both the 1249insG and 7945delA frameshift mutations were detected in two out of eight unrelated IV patients; therefore, a larger cohort of IV patients will be useful in assessing the prevalence of these mutations in Singaporean Chinese IV patients. It will also be of considerable interest to screen Singaporean and Chinese AD case series for the presence of these mutations as well as those previously shown to be prevalent in other Asian populations (Nomura *et al.*, 2007; Sandilands *et al.*, 2007); however, we do not currently have access to such a study cohort.

Mutation E2422X was first reported in a Dutch IV patient (Sandilands *et al.*, 2007). It is intriguing to find the identical mutation in a Singaporean Chinese patient as well as a single case in a Chinese control individual. As this mutation was not detected in other European populations nor was it recurrent in the Dutch, it was classified as a rare mutation. Upon further investigation, we were able to identify Chinese ancestry in the Dutch IV patient. Therefore, our present data

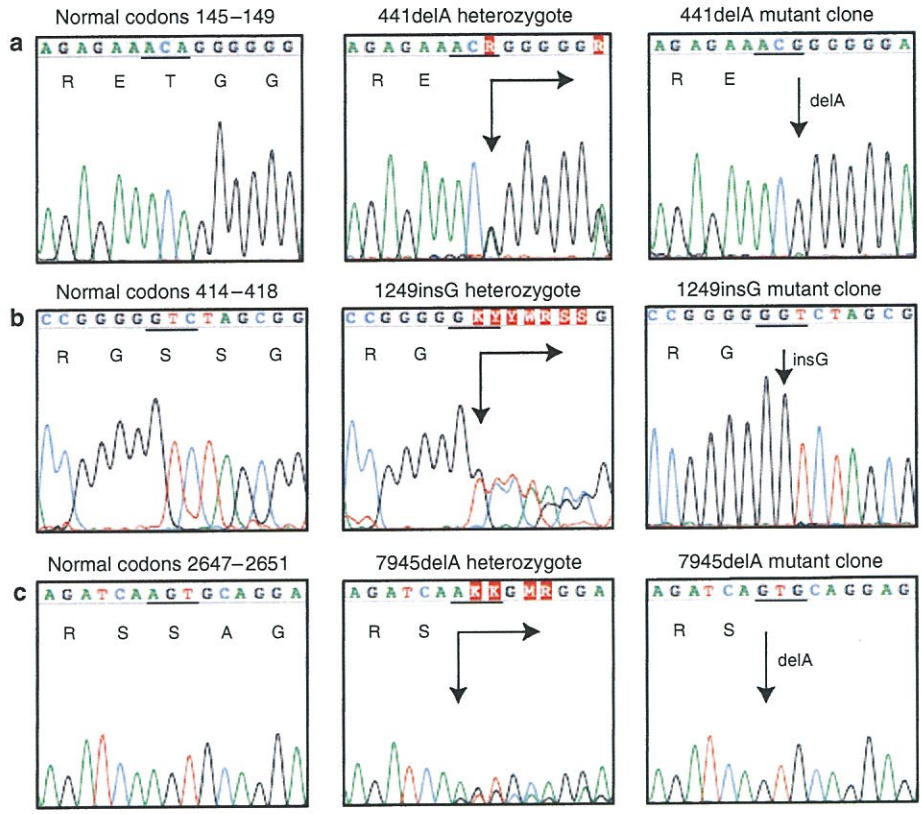



**Figure 3. Frameshift mutations of patients with IV.** (**a**) Direct DNA sequencing of specific *FLG* PCR products. Left panel shows the normal control sequence from the *FLG* S100 domain corresponding to codons 145–149. Middle panel shows the detection of a heterozygous 441delA mutation in patient 8. Right panel confirms the 441delA mutation by cloning and sequencing of the mutant allele. (**b**) Left panel shows the normal control sequence from *FLG* partial repeat ("repeat 0") corresponding to codons 414–418. Middle panel shows the identification of a heterozygous 1249insG mutation detected in patients 6 and 8. Right panel confirms the 1249insG mutation by cloning and sequencing of the mutant allele. (**c**) Left panel shows the normal control sequence from *FLG* repeat 7 corresponding to codons 2647–2651. Middle panel shows the detection of a heterozygous 7945delA mutation detected in patient 6. Right panel confirms the 7945delA mutation by cloning and sequencing of the mutant allele.

suggests that this mutation may be present at low levels in some Asian populations. Ideally, all detected mutations thus far should be screened in different population samples to achieve a more complete picture of the *FLG* mutations in different ancestral groups, but this can be time consuming unless high-throughput screening methods for each mutation can be developed.

Among the eight IV cases studied, patients 5 and 8 are compound heterozygotes for *FLG* mutations and immunohistochemical analysis of skin biopsies showed filaggrin to be nearly absent in the granular layer, whereas patients 2, 3, 4, and 6 who are heterozygous for *FLG* mutations showed reduced filaggrin staining compared with the control (Figure 2). In view of this, it is clearly useful to stain the skin biopsies, when available, for filaggrin to predict the likely number of filaggrin mutations carried by an IV patient. However, it has become apparent that all premature termination codon mutations in *FLG* lead to destabilization of profilaggrin, as observed with human mutations (Sandilands *et al.*, 2007). Thus, it is not possible to carry out precise quantification of the amount of filaggrin repeats in the epidermis of a patient and use this information as a predictive tool to pinpoint the location of mutations along the *FLG* gene.

Immunohistochemical studies of skin biopsy from patient 1 revealed normal levels of filaggrin; therefore, it was not surprising that no *FLG* mutations were detected. As IV and X-linked ichthyosis share similar clinical manifestations (Cuevas-Covarrubias *et al.*, 1998) and this patient is male, we considered the possibility that this individual might harbor a mutation in the steroid sulfatase gene (Shapiro *et al.*, 1978). Subsequently, all 10 exons of the steroid sulfatase gene were fully sequenced but no mutation was identified, thus excluding X-linked ichthyosis. Upon further investigation of the family history, the patient's daughter was found to be similarly affected, suggestive of dominant inheritance. This family therefore appears to have a form of ichthyosis clinically similar to IV, but where immunohistochemical and molecular analysis tend to exclude a filaggrin defect.

Patient 7 exhibited a severe IV phenotype and showed complete absence of filaggrin staining, but no mutation was detected in the exon 3. Further re-sequencing also confirmed that the patient had no detectable mutations in exons 1, 2, or

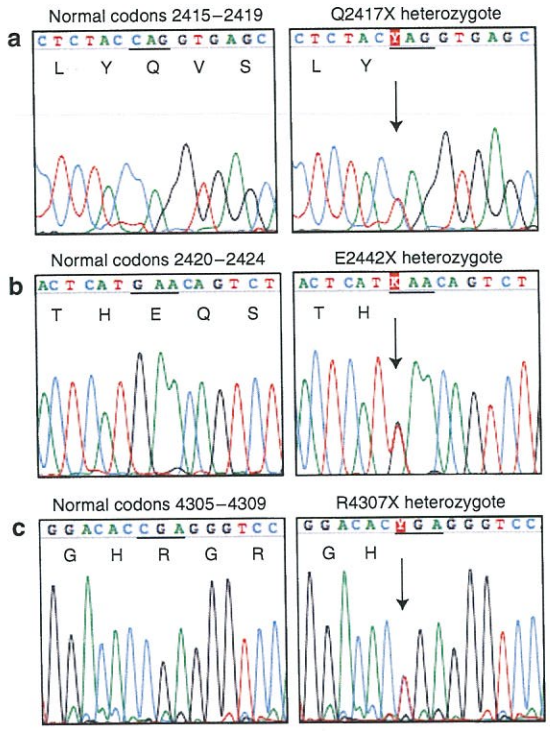



**Figure 4. Nonsense mutations of patients with IV.** (**a**) Direct DNA sequencing of specific *FLG* PCR products. Left panel shows the normal control sequence from the *FLG* repeat 6 corresponding to codons 2415–2419. Right panel shows a heterozygous transition 7249C>T resulting in nonsense mutation Q2147X in patient 6. (**b**) Left panel shows the normal control sequence from *FLG* repeat 6 corresponding to codons 2420–2424. Right panel shows a heterozygous transversion 7264G>T resulting in nonsense mutation E2422X in patient 4. (**c**) Left panel shows the normal control sequence from *FLG* repeat 10.2 corresponding to codons 4305–4309. Right panel shows a heterozygous transition 12919C>T resulting in nonsense mutation R4307X in patient 2.

the splice sites. Thus, it is possible that patient 7 has a mutation in another unknown gene that plays a critical role in the regulation of filaggrin expression, although we also feel it is possible that a mutation has been missed due to the size and complexity of this repetitive gene.

Our study brings the total number of *FLG* mutations reported to date to twenty. The reason why *FLG* mutations are so prevalent is unclear although it has been proposed that balancing selection conferring a heterozygote advantage might explain the persistence and independent emergence of *FLG*-null mutations in different populations (Nomura *et al.*, 2007; Sandilands *et al.*, 2007). It has also been speculated that in individuals carrying one or more *FLG*-null mutation, "natural vaccination" may have permitted the better survival of such heterozygotes during ancient pandemics, due to a stronger background immunity built up by increased trans-epidermal transfer of pathogenic antigens (Irvine and McLean, 2006). Further work on animal models carrying *FLG*-null mutations will be essential to validate this point.

With the emergence of so many different *FLG*-null mutations, more genetic studies on larger numbers of patients will allow us to identify mutation hotspots. Some mutations may also deserve closer examination to study any possible effects on disease penetrance. Current studies report no difference in phenotype between IV patients with *FLG* mutations located in different repeats, but decreased penetrance in AD susceptibility has been reported for certain *FLG*-null mutations (Sandilands *et al.*, 2007). Further *FLG* mutation analysis in different populations will allow us to uncover the full spectrum of mutations and assess their contributions to skin disease and healthcare globally.

## MATERIALS AND METHODS

### Clinical material

Blood samples were obtained from eight independent Singaporean Chinese patients diagnosed with IV and genomic DNA was isolated according to standard procedures. Skin biopsies (3-mm) from the patients were also embedded in formalin and sectioned for immunohistochemistry analysis. In all cases, we followed the Singapore Institutional Review Board/Domain Specific Review Board guidelines, which are in full accordance with the Declaration of Helsinki Principles, and obtained informed written consent from the participants or their parents where the patients were below 21 years of age. DNA from 100 Singaporean and 164 Chinese anonymous, unselected population controls were also included in the study.

### Filaggrin genotyping

*FLG* mutation analysis and allele size variant determination by comprehensive screening of *FLG* in these IV patients was performed using primers and conditions that have been previously reported (Sandilands *et al.*, 2007). Further specific details are available on request. All PCR products were sequenced using an ABI PRISM 3730 genetic analyzer (Applied Biosystems, Foster City, CA). Mutation E2422X was also screened by restriction digestion as reported previously (Sandilands *et al.*, 2007). Previously unreported mutations were confirmed with additional screening methods.

### High-throughput screening for 441delA, 1249insG, and 7945delA

To enable high-throughput screening for the three mutations on control samples, a multiplex fluorescent PCR was used. 50 ng of DNA was amplified in a 10-μl reaction with 1.5 mM $MgCl_2$, 0.25 mM of each dNTP, 5 pmol forward primer, 5 pmol reverse primer, 4% (v/v) DMSO, and 0.5 U AmpliTaq Gold polymerase (Applied Biosystems). The primers used to screen for each mutation and PCR conditions are listed in Table 1. The combined PCR products were diluted 1:60 and sized against ROX-500 size markers according to the manufacturer's recommended protocol (Applied Biosystems). Expected sizes for wild-type and mutant alleles are listed in Table 1.

### Mutation screening for Q2417X

The Q2417X mutation creates an additional *Bfa*I restriction. A 185-bp PCR product was amplified in a 25-μl reaction containing 1.5 mM $MgCl_2$, 10 pmol forward primer 5′-CCACACGTGGCCGGTCAGCA-3′, 10 pmol reverse primer 5′-GTCCTGACCCTCTTGGGACGT-3′, 0.25 mM of each dNTP, 4% (v/v) DMSO, and 1 U high-fidelity polymerase mix (Roche, Penzberg, Germany). PCR amplification conditions were as follows: one cycle of 94 °C (5 minutes) followed by 35 cycles of 94 °C (30 seconds), 64 °C (30 seconds), 72 °C (45 seconds),

**Table 1. Primers and conditions for *FLG* mutation screening by fluorescent PCR**

| Mutation | Primer pairs | Label | Annealing temperature (°C) | Wild-type allele (bp) | Mutant allele (bp) |
|---|---|---|---|---|---|
| 441delA | Forward 5′-GTTTCTTGCTGATAATGTGATTCTGTCT-3′<br>Reverse 5′-ACTAGATTCATGCCTTTTCCC-3′ | HEX | 58 | 390 | 389 |
| 1249insG | Forward 5′-GTTTCTTGGGCAAGCTTCATCTGCAGTC-3′<br>Reverse 5′-CTTGGTGGCTCTGCTGTCTCA-3′ | FAM | 62 | 156 | 157 |
| 7945delA | Forward 5′-GTTTCTTCCCAGGACAAGCAGGAACT-3′<br>Reverse 5′-GTCTTCTGAATGTCCCTCAT-3′ | FAM | 60 | 441 | 440 |

*FLG*, filaggrin gene.
Note: The sequence GTTTCTT was added to the unlabeled forward primer to promote plus-A addition for easier allele calling.

**Table 2. *FLG* profiles of eight Singaporean Chinese IV patients**

| Patient | IV severity | Staining of filaggrin repeats | Mutation(s) detected | Repeat Number | Ethnicity of other cases where found |
|---|---|---|---|---|---|
| 1 | Severe | Normal | — | — | NA |
| 2 | Severe | Reduced | R4307X | 10.2 | — |
| 3 | Mild | Reduced | 7945delA | 7 | — |
| 4 | Mild | Reduced | E2422X | 6 | Dutch, Chinese |
| 5 | Severe | Nearly absent | 1249insG,<br>7945delA | 0, 7 | — |
| 6 | Severe | Reduced | Q2417X | 6 | — |
| 7 | Severe | Absent | — | — | NA |
| 8 | Severe | Nearly absent | 441delA,<br>1249insG | S100 domain,<br>0 | — |

*FLG*, filaggrin gene; IV, ichthyosis vulgaris; NA, not available.

and a final extension cycle of 72 °C (5 minutes). A 5-μl volume of PCR product was incubated with 2.5 U of *Bfa*I (New England Biolabs, Ipswich, MA) in a 20-ml reaction volume at 37 °C overnight and resolved on 3% (w/v) agarose gels. The wild-type allele resolved as an uncut 185-bp fragment, whereas the mutant allele gave fragments of 132 and 53 bp.

### Mutation detection and screening for R4307X

100 ng of DNA was amplified in a 25-μl reaction containing 1.5 mM $MgCl_2$, 10 pmol forward primer 5′-CTCATCATGCAGAGAATTCCT CTG-3′, 10 pmol reverse primer 5′-CTCCAGTACTGGGCCCAGC-3′, 0.25 mM of each dNTP, 4% (v/v) DMSO, and 1 U of high-fidelity enzyme (Roche). PCR amplification conditions were as follows: one cycle of 94 °C (5 minutes) followed by 35 cycles of 94 °C (30 seconds), 58 °C (30 seconds), 72 °C (45 seconds), and a final extension cycle of 72 °C (5 minutes). Mutation R4307X creates an additional *Dde*I site; 5 μl of PCR product was incubated with 2.5 U of *Dde*I (New England Biolabs) in a 20-μl reaction volume at 37 °C overnight and resolved on 3% (w/v) agarose gels. The wild-type allele resolved as 172, 140, 102, and 36-bp fragments, whereas the mutant allele gave fragments of 172, 140, 51, 51, and 36 bp.

### Steroid sulfatase mutation analysis

All exons of the steroid sulfatase gene were amplified by PCR and directly sequenced using the previously reported methods (Liao *et al.*, 2007).

### Immunohistochemistry

Immunoperoxidase staining of paraffin-embedded sections with the Envision system (DakoCytomation, Hamburg, Germany) used the mouse monoclonal 15C10 antibody (Novocastra, Newcastle, UK), which binds to an epitope in the C-terminal portion of the human filaggrin repeat unit. Antigen retrieval was performed by heating sections under pressure for 10–15 minutes in 10 mmol $l^{-1}$ citrate buffer, pH 6. Normal skin samples were stained as controls.

## CONFLICT OF INTEREST

Irwin McLean has filed patents relating to genetic testing and therapy development aimed at the *FLG*.

## ACKNOWLEDGMENTS

We thank the affected individuals for their participation, which made this research possible. We thank the Agency for Science, Technology and Research (A*STAR) Biopolis Shared Facilities (BSF) Histology Unit and DNA Sequencing Unit for pathology and sequencing support, respectively. We thank the following from Ninewells Hospital and Medical School: A Cassidy, G Scott and G McGregor (DNA Analysis Facility) for DNA sequencing and genotyping support; A Grant and A Evans (Pathology Department, Tayside University Hospitals NHS Trust, Dundee) for immunohistochemistry support; and J Mcfarlane (Human Genetics Unit, Dundee) for clerical assistance. Filaggrin research in the McLean laboratory is supported by grants from The British Skin Foundation; The National Eczema Society; The Medical Research Council (reference number G0700314); and donations from anonymous families affected by eczema in the Tayside Region of Scotland. Research in the Institute of Medical Biology is funded by the Biomedical Research Council, A*STAR, Singapore.

## SUPPLEMENTARY MATERIAL

**Figure S1.** Hyperlinearity observed in IV patients at low magnification.

## REFERENCES

Barker JN, Palmer CN, Zhao Y, Liao H, Hull PR, Lee SP *et al.* (2007) Null mutations in the filaggrin gene (FLG) determine major susceptibility to early-onset atopic dermatitis that persists into adulthood. *J Invest Dermatol* 127:564–7

Cuevas-Covarrubias SA, Valdes-Flores M, Orozco Orozco E, Diaz-Zagoya JC, Kofman-Alfaro SH (1998) Are atopy and palm-sole hyperlinearity clinical tools in the differential diagnosis between ichthyosis vulgaris and X-linked ichthyosis? *J Dermatol* 25:556–7

Gruber R, Janecke A, Fauth C, Utermann G, Fritsch P, Schmuth M (2006) Filaggrin mutations p.R501X and c.2282del4 in ichthyosis vulgaris. *Eur J Hum Genet* 15:179–84

Hannifin J, Rajka G (1980) Diagnostic features of atopic dermatitis. *Acta Derm Venereol* 92:44–7

Irvine A (2007) Fleshing out filaggrin phenotypes. *J Invest Dermatol* 127: 504–7

Irvine A, McLean W (2006) Breaking the (un)sound barrier: filaggrin is a major gene for atopic dermatitis. *J Invest Dermatol* 126:1200–2

Kuokkanen K (1969) Ichthyosis vulgaris. A clinical and histopathological study of patients and their close relatives in the autosomal dominant and sex-linked forms of the disease. *Acta Dermato-Venereologica Supplementum* 62:1–72

Liao H, Waters AJ, Goudie DR, Aitken DA, Graham G, Smith FJ *et al.* (2007) Filaggrin mutations are genetic modifying factors exacerbating X-linked ichthyosis. *J Invest Dermatol* 127:2795–8

Marenholz I, Nickel R, Rüschendorf F, Schulz F, Esparza-Gordillo J, Kerscher T *et al.* (2006) Filaggrin loss-of-function mutations predispose to phenotypes involved in the atopic march. *J Allergy Clin Immunol* 118: 866–71

Mischke D, Korge B, Marenholz I, Volz A, Ziegler A (1996) Genes encoding structural proteins of epidermal cornification and S100 calcium-binding proteins form a gene complex ("epidermal differentiation complex") on human chromosome 1q21. *J Invest Dermatol* 106:989–92

Nomura T, Sandilands A, Akiyama M, Liao H, Evans A, Sakai K *et al.* (2007) Unique mutations in the filaggrin gene in Japanese patients with ichthyosis vulgaris and atopic dermatitis. *J Allergy Clin Immunol* 119: 434–40

Palmer C, Irvine A, Terron-Kwiatkowski A, Zhao Y, Liao H, Lee S *et al.* (2006) Common loss-of-function variants of the epidermal barrier protein filaggrin are a major predisposing factor for atopic dermatitis. *Nat Genet* 38:441–6

Presland RB, Haydock PV, Fleckman P, Nirunsuksiri W, Dale BA (1992) Characterization of the human epidermal profilaggrin gene. Genomic organization and identification of an S-100-like calcium binding domain at the amino terminus. *J Biol Chem* 267:23772–81

Rawlings AV, Harding CR (2004) Moisturization and skin barrier function. *Dermatol Ther* 17(Suppl 1):43–8

Sandilands A, Terron-Kwiatkowski A, Hull P, O'regan G, Clayton T, Watson R *et al.* (2007) Comprehensive analysis of the gene encoding filaggrin uncovers prevalent and rare mutations in ichthyosis vulgaris and atopic eczema. *Nat Genet* 39:650–4

Shapiro LJ, Weiss R, Buxman MM, Vidgoff J, Dimond RL, Roller JA *et al.* (1978) Enzymatic basis of typical X-linked icthyosis. *Lancet* 2:756–7

Smith F, Irvine A, Terron-Kwiatkowski A, Sandilands A, Campbell L, Zhao Y *et al.* (2006) Loss-of-function mutations in the gene encoding filaggrin cause ichthyosis vulgaris. *Nat Genet* 38:337–42

Stemmler S, Parwez Q, Petrasch-Parwez E, Epplen JT, Hoffjan S (2007) Two common loss-of-function mutations within the filaggrin gene predispose for early onset of atopic dermatitis. *J Invest Dermatol* 127:722–4

Sybert VP, Dale BA, Holbrook KA (1985) Ichthyosis vulgaris: identification of a defect in synthesis of filaggrin correlated with an absence of keratohyaline granules. *J Invest Dermatol* 84:191–4

Tay YK, Kong KH, Khoo L, Goh CL, Giam YC (2002) The prevalence and descriptive epidemiology of atopic dermatitis in Singapore school children. *Br J Dermatol* 146:101–6

Wells RS (1966) Ichthyosis. *Br Med J* 2:1504–6

HUMAN MUTATION 29(3), 351–360, 2008

*DATABASES*

# The Human Intermediate Filament Database: Comprehensive Information on a Gene Family Involved in Many Human Diseases

Ildiko Szeverenyi,[1] Andrew J. Cassidy,[2] Cheuk Wang Chung,[3] Bernett T.K. Lee,[3] John E.A. Common,[1] Stephen C. Ogg,[1] Huijia Chen,[1] Shu Yin Sim,[1] Walter L.P. Goh,[1] Kee Woei Ng,[1] John A. Simpson,[4] Li Lian Chee,[1] Goi Hui Eng,[1] Bin Li,[1] Declan P. Lunny,[1] Danny Chuon,[3] Aparna Venkatesh,[1] Kian Hoe Khoo,[1] W.H. Irwin McLean,[2] Yun Ping Lim,[3] and E. Birgitte Lane[1,4]*

[1]*Epithelial Biology Group, Institute of Medical Biology, Singapore;* [2]*Epithelial Genetics Group, Human Genetics Unit, Division of Pathology and Neuroscience, University of Dundee, Ninewells Hospital and Medical School, Dundee, United Kingdom;* [3]*Bioinformatics Institute, Singapore;* [4]*College of Life Sciences, University of Dundee, Dundee, United Kingdom*

*Communicated by A. Jamie Cuticchia*

We describe a revised and expanded database on human intermediate filament proteins, a major component of the eukaryotic cytoskeleton. The family of 70 intermediate filament genes (including those encoding keratins, desmins, and lamins) is now known to be associated with a wide range of diverse diseases, at least 72 distinct human pathologies, including skin blistering, muscular dystrophy, cardiomyopathy, premature aging syndromes, neurodegenerative disorders, and cataract. To date, the database catalogs 1,274 manually-curated pathogenic sequence variants and 170 allelic variants in intermediate filament genes from over 459 peer-reviewed research articles. Unrelated cases were collected from all of the six sequence homology groups and the sequence variations were described at cDNA and protein levels with links to the related diseases and reference articles. The mutations and polymorphisms are presented in parallel with data on protein structure, gene, and chromosomal location and basic information on associated diseases. Detailed statistics relating to the variants records in the database are displayed by homology group, mutation type, affected domain, associated diseases, and nucleic and amino acid substitutions. Multiple sequence alignment algorithms can be run from queries to determine DNA or protein sequence conservation. Literature sources can be interrogated within the database and external links are provided to public databases. The database is freely and publicly accessible online at www.interfil.org (last accessed 13 September 2007). Users can query the database by various keywords and the search results can be downloaded. It is anticipated that the Human Intermediate Filament Database (HIFD) will provide a useful resource to study human genome variations for basic scientists, clinicians, and students alike. Hum Mutat 29(3), 351–360, 2008. 

KEY WORDS: database; intermediate filament; keratin; desmin; lamin; GFAP; neurofilament; epidermolysis bullosa; laminopathy

## INTRODUCTION

Intermediate filaments are part of the eukaryotic cell cytoskeleton beside actin filaments and microtubules. The function of the flexible filament proteins appears to be predominantly to provide physical reinforcement for cells in tissues to resist normal wear and tear [Fuchs and Weber, 1994]. Intermediate filament proteins are encoded by 70 genes in the human genome [Hesse et al., 2001; Rogers et al., 2005, 2004], which are divided into six sequence homology groups (types I–VI). Type V filaments, the lamins, are exclusively nuclear and occur in all tissues, but all the others are cytoplasmic filaments showing distinct tissue restriction. Types I and II are the keratins (in epithelial cells), type III are mostly mesodermal (desmin in muscle cells, vimentin in mesenchymal, endothelial, and hemopoietic cells and others; glial fibrillary acidic protein [GFAP] in astroglial cells, and peripherin in subsets of neuronal cells). Type IV are neurofilament and related proteins (NF-L, NF-M, NF-H, nestin, synemin, syncoilin, α-internexin) and type VI are the beaded filament proteins of the eye lens (filensin and CP49).

All of the intermediate filament proteins share a central alpha-helical structure of a long rod domain of highly conserved size and

Received 4 June 2007; accepted revised manuscript 14 August 2007.

*Correspondence to: E. Birgitte Lane, Professor, Institute of Medical Biology, 8A Biomedical Grove, #06-06 Immunos, Singapore 138665. E-mail: birgit.lane@imb.a-star.edu.sg

Grant sponsors: Agency for Science, Technology and Research (A*STAR), Dystrophic Epidermolysis Research Association (DebRA); Grant sponsor: Wellcome Trust; Grant number: 055090/A/98/Z; Grant sponsor: Cancer Research UK; Grant number: C26/A1461.

DOI 10.1002/humu.20652

Published online 21 November 2007 in Wiley InterScience (www.interscience.wiley.com).

variable sized, less structured head (amino-terminal) and tail (carboxy-terminal) domains [Fuchs and Weber, 1994; Steinert and Parry, 1985]. The rod domain has a conserved sequence pattern of heptad repeats (with predominantly hydrophobic residues at the first and fourth position of every heptad) and surface patches of alternating charge, and these features drive the polypeptide chains into a coiled-coil dimer assembly, and thence rapidly [Herrmann and Aebi, 2004; Strelkov et al., 2002] into multistrand filaments of approximately 10 nm wide. Expression of intermediate filament gene transcripts and the small number of their derived alternative splice products is highly tissue-specific, such that intermediate filament proteins have been valuable tools in diagnostic pathology for many years [Domagala et al., 1986; Osborn et al., 1986; Ramaekers et al., 1985; reviewed by Fuchs and Cleveland, 1998; Porter and Lane, 2003]. From studies on the human diseases it is now understood that intermediate filaments are essential for tissue function, as they provide cells with necessary resilience to maintain tissue structure [reviewed by DePianto and Coulombe, 2004].

The number of diseases associated with intermediate filaments is still growing (currently 72), with the first human diseases linked to mutation in genes encoding lamin B1 (*LMNB1*) and filensin (*BFSP1*) reported recently [Padiath et al., 2007; Ramachandran et al., 2007]. The number of intermediate filament mutations reported is expanding even faster [Burke and Stewart, 2006; Capell and Collins, 2006; Goldfarb et al., 2004; Lane and McLean, 2004; Lariviere and Julien, 2004; Omary et al., 2004]. This disease association has generated increased interest worldwide in the protein family, and, consequently, an increasing amount of data on these genes and proteins is appearing in the literature at an accelerating rate. Adding to this the recently published revised gene nomenclature of the largest group, the keratins [Schweizer et al., 2006], it is now timely to combine all the information on these proteins into one easily accessible source. Based on the prototype created 5 years ago at the University of Dundee, we present here a significantly expanded and updated database, the Human Intermediate Filament Database (HIFD), as a useful reference and analysis tool.

## DATA SOURCES AND LIMITATIONS

The HIFD contains cataloged and manually curated data retrieved from publications describing sequence variants (mutations) and allelic variants (polymorphisms) identified in the genes encoding intermediate filaments in humans (Fig. 1). Over 460 peer-reviewed research articles have been mined to date, published between 1991 and 2007, extracted from PubMed (www.ncbi.nlm.nih.gov/entrez) by bibliographic searches using keywords such as *intermediate filament*, *mutation*, *polymorphism*, *variant*, *disease*, *keratin*, *epidermolysis bullosa*, *laminopathy*, *desmin*, and *neurofilament*, to identify primary data describing naturally-arising disease-associated human intermediate filament mutations. Reviews, and studies presenting experimentally induced mutation(s), were not included. The sequence and allelic variants were entered into the database as many times as they were found in unrelated cases. Multiple family members expressing the same variant sequence are only entered as one incidence, in contrast with other databases such as Universal Mutation Database *LMNA* (www.umd.be:2000). Similarly, when a mutation from the same patient has been described in more than one article, only the original description has been taken into account and registered in the database.

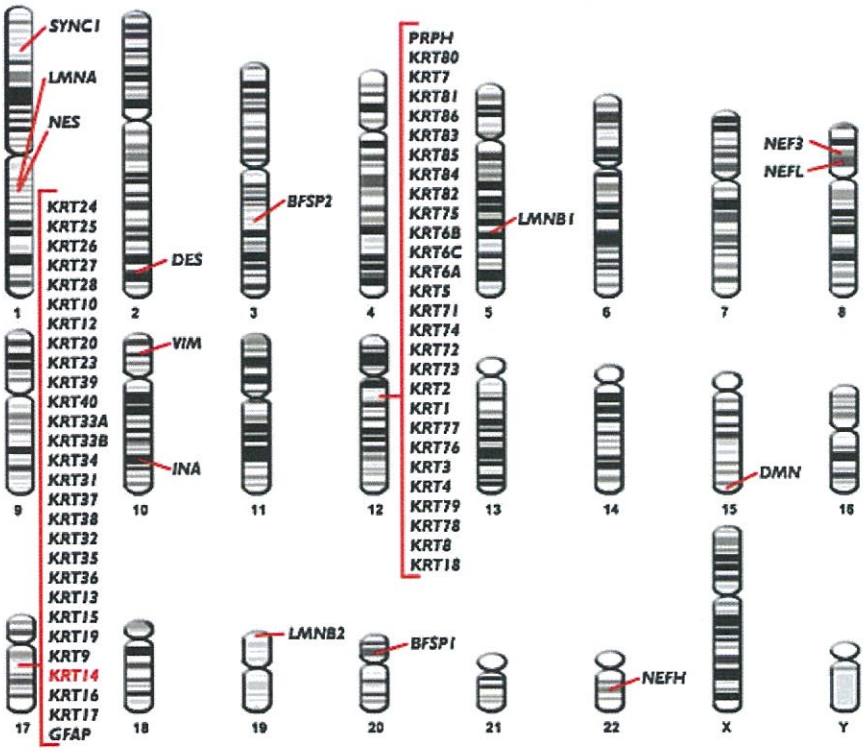



FIGURE 1. **Chromosomal location of human intermediate filament coding genes. Individual protein records can be accessed by clicking on the gene symbols (e.g., *KRT14*) arranged according to their chromosomal locations. Illustration of chromosomes was obtained from Ensembl (www.ensembl.org) and modified.**

## DATABASE CONTENT AND ORGANIZATION

The HIFD is a protein-centric human database listing sequence variants of intermediate filaments associated with various genetically inherited diseases. Known allelic variants (polymorphisms) for intermediate filament genes are also collected and displayed. Data recorded for the 70 intermediate filament genes include information on the primary protein plus the few recorded alternative splice products that are significantly expressed in normal human tissues. This amounts to 73 proteins in total including desmuslin (*DMN*) with two alternative splice products (synemin α and synemin β) and lamin A/C (*LMNA*) with three alternative splice products. The *LMNA* gene encodes three major proteins, lamins A, C1, and C2. Lamin AΔ10 is not included in the database as it has only been detected in human cell lines and not, so far, in normal tissues [Machiels et al., 1996]. At the time of writing there are altogether 1,235 sequence variants and 168 allelic variants in the database. Nearly all of the sequence variants listed are clearly implicated in one or more of 72 distinct human diseases described in the database.

HIFD is a relational database built using MySQL and arranged to facilitate easy access to the relevant information. The data is organized according to types of intermediate filament proteins and their related diseases. Details such as protein name and physicochemical characteristics, gene symbol, genomic location, sequence information, domain positions, and associated diseases are accessible directly through the individual protein information page for each intermediate filament protein. Each sequence variant and allelic variant has a unique database identification number (ID) and they are presented in the form of separate tables. Variations are mapped to cDNA sequences and the changes are described at DNA and protein levels with three-letter amino acid codes. The positions of the variants referred in the papers were curated (and revised if necessary) to fit the consensus nomenclature of the Human Genome Variation Society (www.hgvs.org/mutnomen) [den Dunnen and Antonarakis, 2000, 2001]. Accordingly, the first coding nucleotide position is taken as the adenine of the initiation (ATG) codon in the cDNA reference sequence, and ATG is numbered as the first codon (encoding methionine) of the protein. Following current convention, methionine is taken as the first amino acid of the protein sequence, irrespective of whether it is subsequently cleaved. The variants were assigned to the integrated nucleic and amino acid sequences to provide a comprehensive view of the data set (Fig. 2A).

The genomic, mRNA and protein sequences were retrieved from the NCBI RefSeq (www.ncbi.nlm.nih.gov/projects/RefSeq) collection [Pruitt et al., 2007] and they are accessible for downloading in FASTA (text-based) format, while the variant records are downloadable in Excel (Microsoft, Redmond, WA) format. The short descriptions of diseases are organized according to their association with intermediate filament type and links are included to the respective protein information page, to references in the HIFD, as well to the Online Mendelian Inheritance in Man (OMIM) database. The database statistics section reviews the records for variants according to sequence homology group, variant type, affected domain, associated diseases, and nucleic or amino acid substitution.

In addition to the data repository and visualization components, the HIFD also includes the frequently used CLUSTALW tool [Thompson et al., 1994], which has been incorporated within the web portal to run on our backend computer platforms. This resource allows multiple sequence comparison of orthologous nucleotide and protein sequences as obtained from HomoloGene (Fig. 2B) and can be visualized using the JalView tool [Clamp et al., 2004]. Hyperlinks to public databases such as HGNC (HUGO Gene Nomenclature Committee; www.genenames.org/index.html), OMIM (www.ncbi.nlm.nih.gov/entrez/query.fcgi?db = OMIM), NCBI Entrez Gene (www.ncbi.nlm.nih.gov/entrez/query.fcgi?db = gene), and the University of California, Santa Cruz (UCSC) Genome Browser (UCSC Genome Bioinformatics Group; http://genome.ucsc.edu) are all provided.

Information in HIFD is periodically updated to ensure continuity and accuracy. Contact details and an online feedback form are available for the user community to contribute their own comments and updates. During its "soft-launch" phase, the database has already benefited from suggestions of several local and international users who have already started to use it for their research.

## QUERYING THE DATABASE

The database is freely and publicly accessible online at www.interfil.org. Individual protein records can be accessed either by clicking onto the gene symbols displayed according to their chromosomal location (Fig. 1) or via the navigation bar. The major advantages of HIFD are that it allows the user to search at the different levels of information and the query results can be downloaded as Excel spreadsheets. The database is searchable by several keywords (e.g., protein name, gene symbol, cDNA and protein RefSeq accession numbers, chromosome number/location, associated disease, variant type, and publication records). The predetermined drop-down menus are displayed to help nonspecialist users in data query. Users can submit sequences directly into database to search for similarities using the Basic Local Alignment Search Tool (BLAST) [Altschul et al., 1990]. Data can be retrieved by browsing through either the entire collection of intermediate filaments organized into six different types of homology group or the full bibliography containing curated references listed in alphabetical order and linked to PubMed.

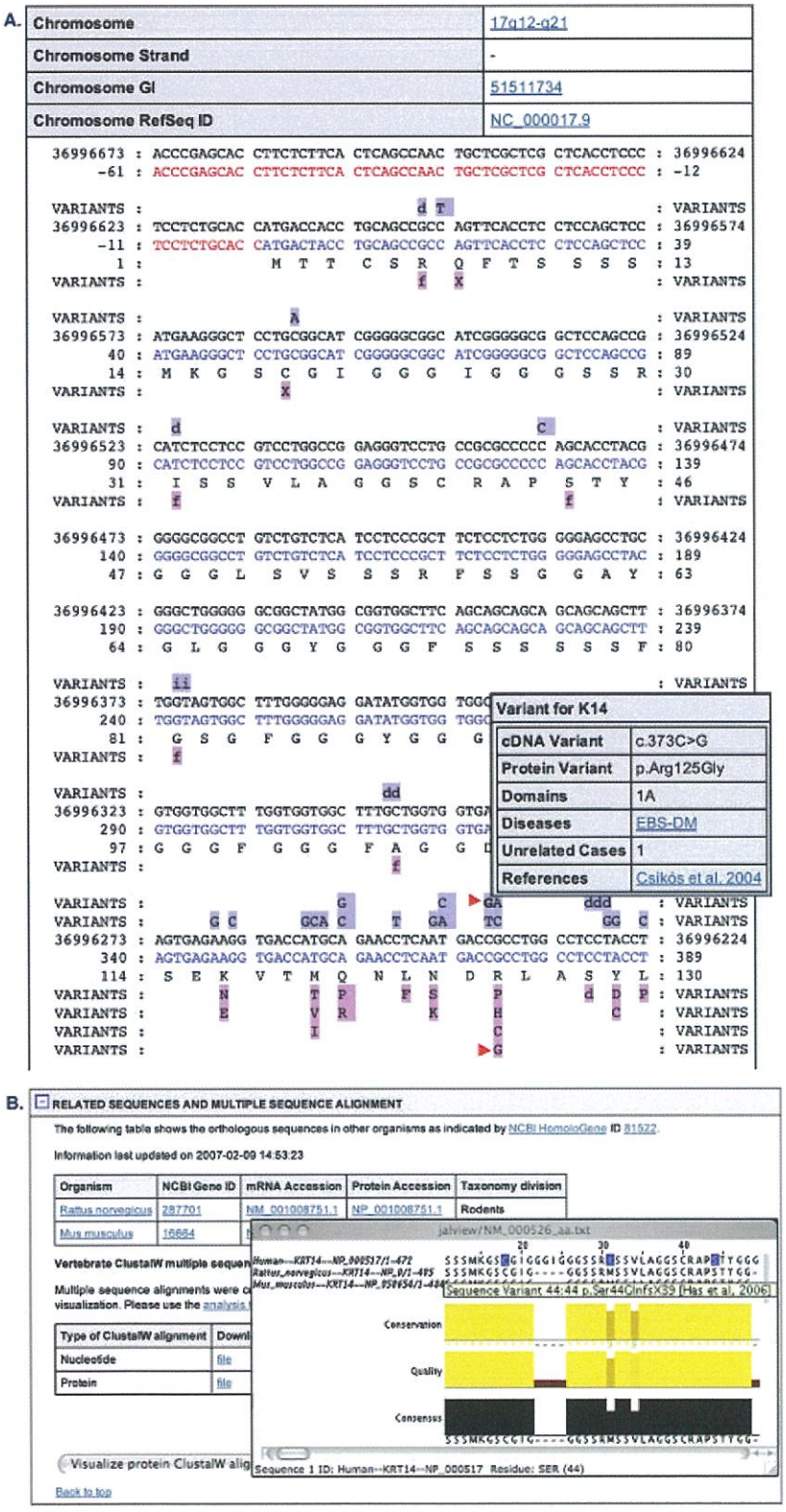
## ANALYSIS OF DATASETS

There are 73 intermediate filament proteins encoded within the human genome [Hesse et al., 2001; Rogers et al., 2005, 2004], and 34 of these have so far been associated with diseases in humans. Some of the proteins have been linked with multiple diseases, leading to a total of 72 intermediate filament-associated diseases in the database (Fig. 3).

To date, the single intermediate filament protein with the highest number of recorded unrelated cases of pathogenic mutations is lamin A, with a total of 333 sequence variants; 18 of the 20 described laminopathies are associated with the *LMNA* gene. This is followed by K14 with 152 recorded sequence variants, GFAP with 127, and K5 with 122 sequence variants. Since K5 and K14 are always coexpressed in the same cells as obligatory copolymers, mutations in either K5 or K14 cause the same pathology: combined reports of mutations in K5 and K14 gives the next highest category of known intermediate filament pathologies (274 disease-associated sequence variants) after lamin A.

Of all the intermediate filament proteins known, 48% (35/73) have at least one sequence or allelic variant record. Most of the proteins with no recorded sequence variants are the keratins

associated with the hair follicle, a dense and complex structure which is hard to analyze histologically and whose proteins are very insoluble and hard to distinguish from one another. Only four of the 31 known hair follicle-associated intermediate filament proteins have been linked to diseases. Assuming that the disease-associated K8 and K18 mutations can be classed as pathogenic or pathocontributory, then outside the hair keratins, only *KRT15*, *KRT19*, *KRT7*, *VIM* (vimentin), *INA* (internexin),



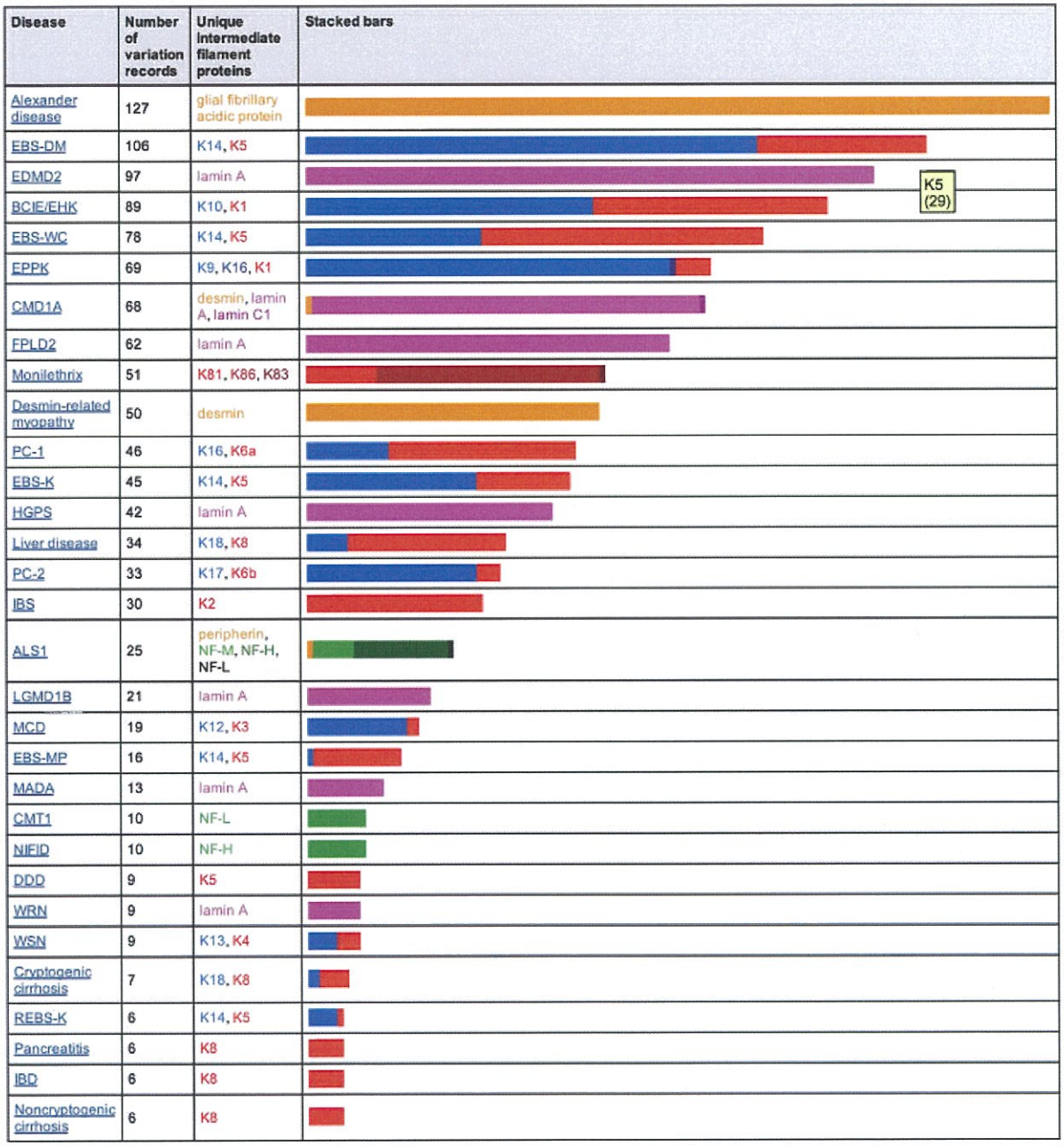



FIGURE 3. Screenshot of the most highly represented human genetic diseases listed in the database (31 shown out of 72 in total). Bars represent numbers of pathogenic sequence variants reported. Proteins are labeled with different colors according to their homology groups. The "mouse-over" feature of the bar chart allows the user to identify the number of variants of the respective proteins associated with each disorder.

FIGURE 2. Sequence alignment interface. **A:** Integrated sequence view. The screenshot captures part of the alignment of genomic, mRNA, and protein sequences of K14. Red and blue colored sequences denote untranslated and coding regions of mRNA, respectively. Nucleic acid variants are displayed in boxed blue background above the integrated view, while changes in the wild-type protein sequences are illustrated in single-letter amino acid code in boxed pink background below the sequence. The pop-up box with information on genomic variation, domains, unrelated patient cases, associated diseases, and references can be displayed by moving the mouse over a DNA or protein variant (red arrowheads). Hyperlinks to OMIM and PubMed are included in the pop-up box. **B:** Orthologous sequence alignment. Shown here is a screenshot of human, mouse, and rat K14 protein sequences, aligned and visualized in Jalview. The pop-up window shows the alignment of orthologous sequences with graphical display of protein conservation and with the consensus sequence below. Sequence variants are labeled with blue background with "mouse-over" option allowing for the display of data entry references.

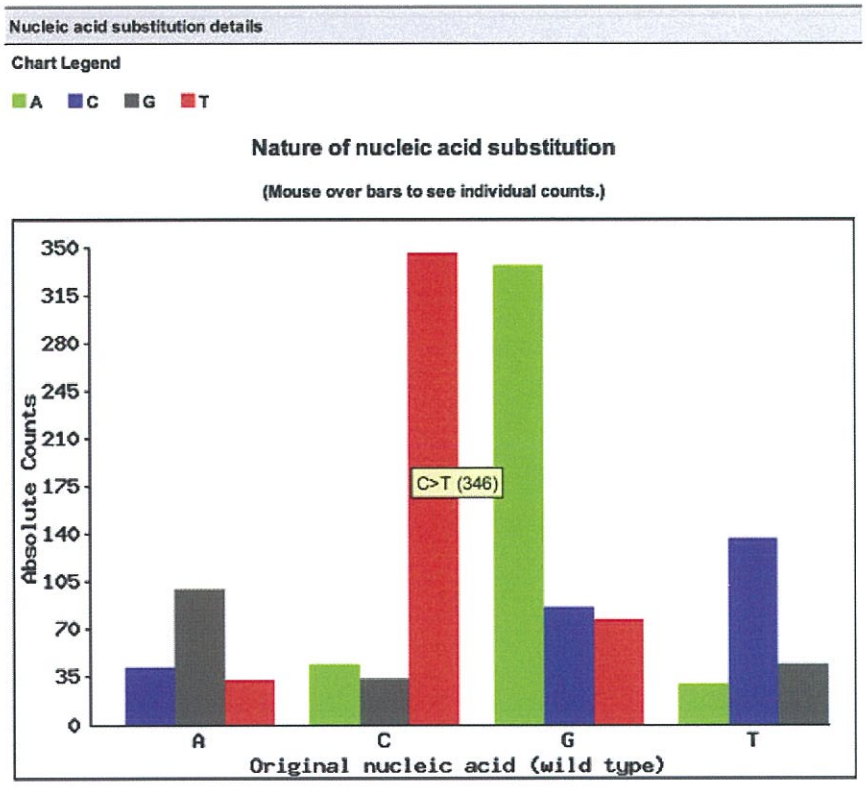



FIGURE 4. **Nucleotide substitution pattern in human intermediate filament genes. The "mouse-over" pop-up box shows the occurrence of a particular point mutation. Disease causing mutations are predominantly C to T and G to A transitions at the DNA level.**

*NES* (nestin), *SYNC1* (syncoilin), and *DMN* (desmuslin) have not been linked to any human disease. It should be suspected that mutation in these genes may be either lethal or clinically asymptomatic.

The database currently contains 1,274 sequence variants and 170 allelic variants collected from 459 peer-reviewed articles. The genomic variations are classified according to the changes in protein sequences: 76.7% missense mutations, 8.4% silent mutations, 5.2% deletions, 3.3% frameshift mutations, 1.9% nonsense mutations, 1.6% silent/deletion (a silent mutation in *LMNA* activates a cryptic splice site resulting in an in-frame deletion), 0.7% insertions, 0.6% insertion-deletions, and 0.07% duplications. There are residual uncertainties (1.4% unknown) when changes at protein level have not been described but several possibilities exist based on the DNA sequence variations.

We analyzed the nature of substitution in our dataset, since this mutation type cause the majority of changes in the genetic information. By far the most frequent substitutions were C to T (C>T) and G to A (G>A); these substitutions are known to be classically associated with methylation events. Cooper and Youssoufian [1988] have described the contribution of methylation localized to the coding region of DNA to human genetic diseases in general. In the case of sequence variants listed in HIFD, transitions from C to T and G to A took place 2.8 times more frequently than in the opposite direction (T to C and A to G) (Fig. 4). The reason for this difference is that the spontaneous deamination of 5-methylcytosine in CpG dinucleotides creates thymine [Duncan and Miller, 1980], which cannot be repaired by uracil-DNA glycolase repair machinery. In contrast, deamination of unmethylated cytosine produces uracil, which is amenable to correction by uracil-DNA glycosylase [Duncan and Miller, 1980; Lindahl et al., 1997].

TABLE 1. **Arginine Mutation Hotspots Within the Homology Groups**

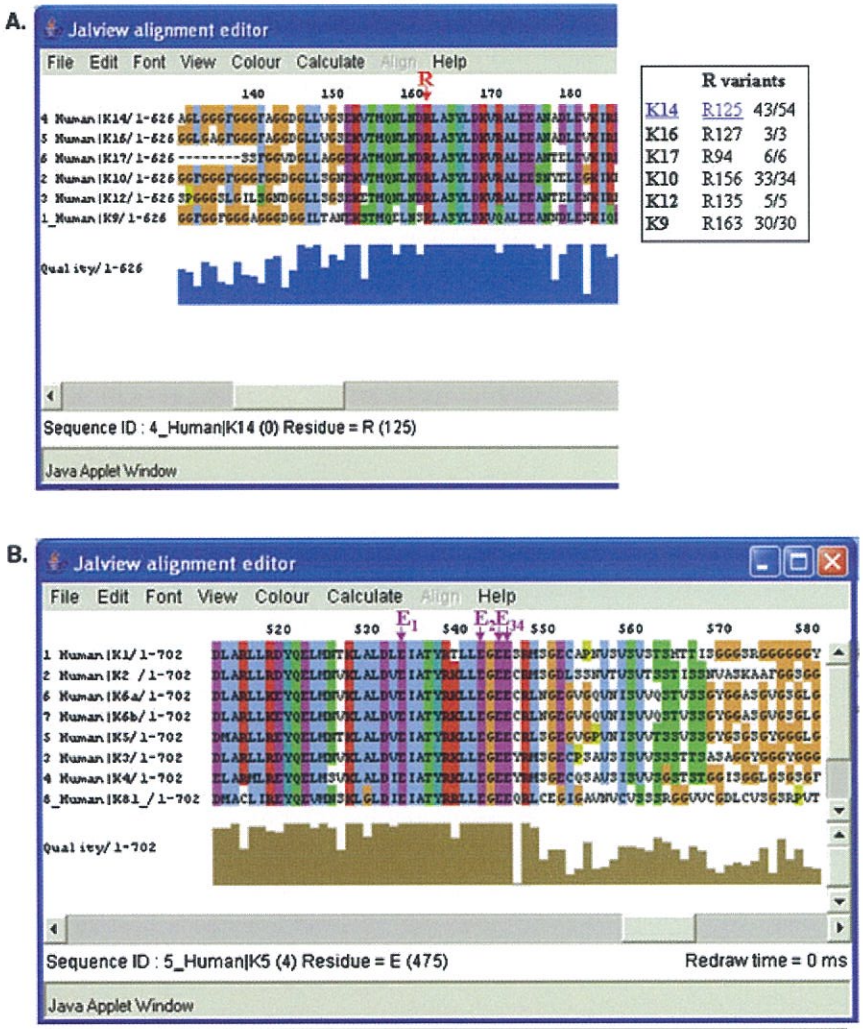
| IF | Arg (R) | K14:R125[a] | K14:R288[a] | K14:R416[a] | lamin A:R482[b] |
|---|---|---|---|---|---|
| Domain Homology group | Total | 1A | 2A | 2B | TAIL |
| Type I | 133 | 122 | 0 | 1 | NA |
| Type II | 23 | 0 | 0 | 0 | NA |
| Type III | 87 | 20 | 30 | 8 | NA |
| Type IV | 1 | 0 | 0 | 0 | 0 |
| Type V | 152 | 0 | 7 | 11 | 48 |
| Type VI | 1 | 0 | 0 | 0 | NA |
| Total | 397 | 142 (36%) | 37 (9%) | 20 (5%) | 48 (12%) |

[a]Number of Arg mutations at corresponding residues of K14.
[b]Number of Arg mutations at corresponding residues of lamin A.
NA, not applicable.

At the protein level, the highest mutation frequency was found in arginine (Arg, R) codons, due to the high mutability of the CpG dinucleotide (see previous paragraph) in four out of six possible arginine codons. Most of the arginine variants were identified in types I, III, and V homology groups (Table 1). In K14, sequence variation of codon 125 (Arg) accounts for 43 out of 54 arginine mutations. This residue lies within one of the most highly conserved regions of type I keratins, the helix initiation motif, and its mutation has been associated the most severe subtype of the skin-blistering disorder epidermolysis bullosa simplex [Lane and

McLean, 2004; Letai et al., 1993]. ClustalW alignment of K14 sequence with other type I protein was used to locate the equivalent arginine to K14: R125. The alignment confirms that this is a mutation hotspot for type I group with all of the sequence variants for arginine in K9, K10, K12, K16, and K17 in the equivalent position to K14: R125 (Fig. 5A).

Alongside type I keratins, lamin A (*LMNA*) and glial fibrillary acidic protein (*GFAP*) have contributed most to the arginine sequence variants catalogued in the database (Table 1). The majority of the mutated Arg residues (100/152) to date are in the tail domain of lamin A, predominantly at position R482, which with 48 reports is a mutation hotspot. Mutation of this residue was




| | R variants | | |
|---|---|---|---|
| K14 | R125 | 43/54 | |
| K16 | R127 | 3/3 | |
| K17 | R94 | 6/6 | |
| K10 | R156 | 33/34 | |
| K12 | R135 | 5/5 | |
| K9 | R163 | 30/30 | |

| | $E_1$ variants | | $E_2$ variants | | $E_3$ variants | | $E_4$ variants | |
|---|---|---|---|---|---|---|---|---|
| K1 | E478 | 2/5 | E487 | - | E489 | 1/5 | E490 | 2/5 |
| K2 | E476 | 2/20 | E485 | - | E487 | 17/20 | E488 | 1/20 |
| K3 | E498 | - | E507 | - | E509 | 1/1 | E510 | - |
| K4 | E509 | - | E518 | - | E520 | - | E521 | - |
| K5 | E466 | - | E475 | 4/18 | E477 | 8/17 | E478 | - |
| K6a | E461 | - | E470 | - | E472 | 1/1 | E473 | - |
| K6b | E461 | - | E470 | - | E472 | 3/3 | E473 | - |
| K81 | E402 | 6/8 | E411 | - | E413 | 2/8 | E414 | - |
| K83 | E407 | 1/1 | E416 | - | E418 | - | E419 | - |
| K86 | E402 | 9/34 | E411 | - | E413 | 25/34 | E414 | - |

FIGURE 5. ClustalW alignment of keratin proteins. These screenshots show the most prominent hotspot regions for pathogenic mutations in keratins, the helix boundary regions. Comparison of the amino acid sequence of (**A**) coil 1A (helix initiation region) of type I keratins and (**B**) coil 2B (helix termination region) of type II keratins. The most frequently mutated amino acids, Arg in type I (R) and Glu in type II ($E_{1-4}$), are located at equivalent positions within each homology group. The relative positions of these amino acids in the corresponding sequences are listed in the adjacent table. The fractional numbers in the tables indicate the frequency with which this mutation occurred in this specific position, out of the total number of mutated residues (Arg or Glu) in that sequence. The reference positions (K14 for type I and K5 for type II) are underlined.

described to cause Dunnigan-type lipodystrophy [Broers et al., 2006; Burke and Stewart, 2006; Cao and Hegele, 2000; Capell and Collins, 2006; Rogers et al., 2005; Shackleton et al., 2000; Speckman et al., 2000; Vigouroux et al., 2000]. Arginine mutation clusters are very common in GFAP: 31 of them are in the 1A domain at codons R79 (equivalent position to K14:R125) and R88 (corresponding to K14:R134), and 30 are in the 2A domain at residue R239 (corresponding to K14:R288). Together, only 12 Arg variants were entered for desmin and eight of them were to be found at R406 (corresponding to K14:R416).

The helix initiation motif at the start of the rod domain interacts directly with the helix termination motif at the end of the rod domain of the next subunit in the filament and has an essential role in filament assembly [Herrmann and Aebi, 1998]. The second most frequently mutated residue in the database was glutamic acid (Glu, E), and the majority of these mutations lie within the helix termination motif (see Table 2). Analysis of the distribution of pathogenic sequence variants of glutamic acid revealed that 92 out of 142 mutations affecting glutamic acid were found in type II keratins in the helix termination motif, while 133 out of 397 of the Arg mutations were in a functionally reciprocal region in type I keratins in the helix initiation motif (Tables 1 and 2). Two functionally significant conserved glutamic acid residues were found in helix termination motif of the type II group (Fig. 5B). The most common mutation in type II keratins is one of these, E477 in K5 (8/17 K5 substitutions) and its equivalent residue in other type II sequences. This position was described as a hotspot for K86 (formerly known as Hb6) in monilethrix patients and for K2 in association to ichthyosis bullosa of Siemens [Korge et al., 1998; Rothnagel et al., 1994]. The second most common Glu sequence variant is 11 residues farther toward the amino terminal, although interestingly this has not been found to be mutated in K5 or K6a/b. The database contains 16 glutamic acid mutations in GFAP; half of them are in the hotspot region equivalent position to K5:E477 and the other half are mostly in the 1B domain. There are 18 glutamic acid mutation records for lamin A, but they are scattered along the protein sequence.

Of the glutamic acid changes, 79% were changed to lysine, leading to altered charge in the protein, which would be predicted to have a destabilizing effect on the filament [Rothnagel et al., 1994]. The equivalent position to K5:466 is highly conserved in charge across the intermediate filament proteins and this is important for early filament assembly and stability [Wu et al., 2000].

It is notable that to date there have been no sequence variants identified in K8 or K18 corresponding to the hotspots in helix 1A (K14: R125 equivalent) or helix 2B (K5: E477 equivalent), presumably because these highly disruptive mutations may be lethal in these keratins that are expressed very early in embryogenesis.

TABLE 2. Glutamic Acid Mutation Hotspots Within the Homology Groups

| IF | Glu (E) | K5:466[a] | K5:475[a] | K5:477[a] | K5:478[a] |
|---|---|---|---|---|---|
| Domain | Total | 2B | 2B | 2B | 2B |
| Homology group | | | | | |
| Type I | 5 | 1 | 0 | 0 | 0 |
| Type II | 92 | 20 | 4 | 58 | 3 |
| Type III | 18 | 1 | 1 | 5 | 2 |
| Type IV | 9 | 0 | 0 | 0 | 2 |
| Type V | 18 | 0 | 0 | 0 | 0 |
| Type VI | 0 | 0 | 0 | 0 | 0 |
| Total | 142 | 22 (15%) | 5 (4%) | 63 (44%) | 7 (5%) |

[a]Number of Glu mutations at corresponding residues of K5.

The majority of the disease-causing variations (primarily in types I–III) are found in the helix boundary motifs of the α-helical subdomains of 1A and 2B (375 and 234 records, respectively), which is in accordance with the highly conserved molecular architecture (i.e., low tolerance of sequence variation) of these regions across intermediate filaments [Strelkov et al., 2002], and the apparent functional importance of these regions for filament assembly. Another potential contributor to this bias, however, could be that end domain mutations in keratins have been systematically underreported, because of a tendency to sequence primarily the two major hotspots. High numbers of mutations in the tail domain have only been recorded in the lamins (200/232): lamins have several important functional motifs in their tail domains, including a nuclear localization signal (NLS), a region with immunoglobulin-like structure and a CaaX motif for membrane anchorage modifications. Sequence variation in the NLS results in aberrant filament assembly in the cytoplasm, and mutations affecting the Ig-like structure or the CaaX-dependent posttranslational processing have been directly associated with laminopathies [Krimm et al., 2002; Loewinger and McKeon, 1988].

Thus it is apparent that there are some very prominent hotspots of mutation, which are seen in some types of intermediate filaments but not all. Reasons for this may lie in the biology of the cell and tissue types in which these proteins are major structural components. Severe or catastrophic mutations, such as those in the rod ends of keratins, may only be compatible with life when they occur in tissues where cells are naturally replaced very frequently (such as in the epidermis).

## DATA SUBMISSION

The growth of the HIFD is envisioned to be a collective effort from the scientific community. Users are encouraged to submit new variants even if they are unpublished, and it is anticipated that this facility will be increasingly used as the database grows. An online submission form is available for user submissions, which is a user-friendly simplified interpretation of the recommendations of the Human Genome Variation Society. Each new variant will be allocated a unique identifier based upon intermediate filament type and date of acceptance, for publication or for unpublished incorporation into the database. For variants submitted to the database but not submitted for publication to a scientific journal, the identifier number will be allocated on the basis of acceptance by a curator committee: a senior member of staff from each of the collaborating research or clinical groups forms part of a committee to review unpublished variants submitted to the Database.

A web-based curation portal has been developed as part of the maintenance effort to enter new records and maintain updated information in the Database. The portal facilitates the curation workflow, allows easy data entry and enforces basic quality control. Access to the curation portal is password protected and is available only to a group of curators.

## CONCLUSION AND FUTURE PROSPECTS

Given the increasing interest toward and growing importance of intermediate filament associated diseases, and the massive amount of related literature published each year, it is becoming progressively harder for any single researcher to cover the research field. The existence of a comprehensive database such as this, the

HIFD, should be a considerable help to all workers in this and related fields. The HIFD is a freely available repository for pathogenic sequence variants and allelic variants of intermediate filament genes, equally useful for clinicians investigating phenotype–genotype correlations, population geneticists studying the distribution of mutations, cell and molecular biologists working on functional analyses, and students studying intermediate filament genes. Moreover, the patterns of mutation incidence that emerge from such a database can inform strategies in clinical genetics, such as, e.g., the necessity to sequence the whole coding region. In addition to the manually curated information, the HIFD integrates data from a variety of resources (e.g., HGNC, NCBI Entrez Gene, PubMed, OMIM, and RefSeq), allowing them to be accessed from a single web portal. We hope that this database succeeds in providing a resource that integrates analysis tools for current data with a user-friendly graphical interface.

It is planned to keep the database current with online data submissions, frequent updates, and a rigorous curation process, to retain and increase its value as a resource for academic, clinical, and commercial users alike. The structure of the database allows for expansion to include further data categories (e.g., a list of antibodies against intermediate filament proteins) and new features, should they be required by the user community.

## ACKNOWLEDGMENTS

We are especially grateful to David A.D. Parry, Massey University, New Zealand, for help with the protein domain annotation and to Michael A. Rogers, Deutsches Krebsforschungszentrum, Heidelberg, Germany, for providing updated reference sequences. We thank Sarina Tay for the web design, the many colleagues who have contributed data to the project, and to users for their feedback and support.

## REFERENCES

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. Basic local alignment search tool. J Mol Biol 215:403–410.

Broers JL, Ramaekers FC, Bonne G, Yaou RB, Hutchison CJ. 2006. Nuclear lamins: laminopathies and their role in premature ageing. Physiol Rev 86:967–1008.

Burke B, Stewart CL. 2006. The laminopathies: the functional architecture of the nucleus and its contribution to disease. Annu Rev Genomics Hum Genet 7:369–405.

Cao H, Hegele RA. 2000. Nuclear lamin A/C R482Q mutation in Canadian kindreds with Dunnigan-type familial partial lipodystrophy. Hum Mol Genet 9:109–112.

Capell BC, Collins FS. 2006. Human laminopathies: nuclei gone genetically awry. Nat Rev Genet 7:940–952.

Clamp M, Cuff J, Searle SM, Barton GJ. 2004. The Jalview Java alignment editor. Bioinformatics 20:426–427.

Cooper DN, Youssoufian H. 1988. The CpG dinucleotide and human genetic disease. Hum Genet 78:151–155.

den Dunnen JT, Antonarakis SE. 2000. Mutation nomenclature extensions and suggestions to describe complex mutations: a discussion. Hum Mutat 15:7–12.

den Dunnen JT, Antonarakis SE. 2001. Nomenclature for the description of human sequence variations. Hum Genet 109:121–124.

DePianto D, Coulombe PA. 2004. Intermediate filaments and tissue repair. Exp Cell Res 301:68–76.

Domagala W, Lubinski J, Weber K, Osborn M. 1986. Intermediate filament typing of tumor cells in fine needle aspirates by means of monoclonal antibodies. Acta Cytol 30:214–224.

Duncan BK, Miller JH. 1980. Mutagenic deamination of cytosine residues in DNA. Nature 287:560–561.

Fuchs E, Cleveland DW. 1998. A structural scaffolding of intermediate filaments in health and disease. Science 279:514–519.

Fuchs E, Weber K. 1994. Intermediate filaments: structure, dynamics, function, and disease. Annu Rev Biochem 63:345–382.

Goldfarb LG, Vicart P, Goebel HH, Dalakas MC. 2004. Desmin myopathy. Brain 127(Pt 4):723–734.

Herrmann H, Aebi U. 1998. Structure, assembly, and dynamics of intermediate filaments. Subcell Biochem 31:319–362.

Herrmann H, Aebi U. 2004. Intermediate filaments: molecular structure, assembly mechanism, and integration into functionally distinct intracellular Scaffolds. Annu Rev Biochem 73:749–789.

Hesse M, Magin TM, Weber K. 2001. Genes for intermediate filament proteins and the draft sequence of the human genome: novel keratin genes and a surprisingly high number of pseudogenes related to keratin genes 8 and 18. J Cell Sci 114(Pt 14):2569–2575.

Korge BP, Healy E, Munro CS, Pünter C, Birch-Machin M, Holmes SC, Darlington S, Hamm H, Messenger AG, Rees JL, Traupe H. 1998. A mutational hotspot in the 2B domain of human hair basic keratin 6 (hHb6) in monilethrix patients. J Invest Dermatol 111:896–899.

Krimm I, Ostlund C, Gilquin B, Couprie J, Hossenlopp P, Mornon JP, Bonne G, Courvalin JC, Worman HJ, Zinn-Justin S. 2002. The Ig-like structure of the C-terminal domain of lamin A/C, mutated in muscular dystrophies, cardiomyopathy, and partial lipodystrophy. Structure 10:811–823.

Lane EB, McLean WH. 2004. Keratins and skin disorders. J Pathol 204:355–366.

Lariviere RC, Julien JP. 2004. Functions of intermediate filaments in neuronal development and disease. J Neurobiol 58:131–148.

Letai A, Coulombe PA, McCormick MB, Yu QC, Hutton E, Fuchs E. 1993. Disease severity correlates with position of keratin point mutations in patients with epidermolysis bullosa simplex. Proc Natl Acad Sci USA 90:3197–3201.

Lindahl T, Karran P, Wood RD. 1997. DNA excision repair pathways. Curr Opin Genet Dev 7:158–169.

Loewinger L, McKeon F. 1988. Mutations in the nuclear lamin proteins resulting in their aberrant assembly in the cytoplasm. EMBO J 7:2301–2309.

Machiels BM, Zorenc AH, Endert JM, Kuijpers HJ, van Eys GJ, Ramaekers FC, Broers JL. 1996. An alternative splicing product of the lamin A/C gene lacks exon 10. J Biol Chem 271:9249–9253.

Omary MB, Coulombe PA, McLean WH. 2004. Intermediate filament proteins and their associated diseases. N Engl J Med 351:2087–2100.

Osborn M, van Lessen G, Weber K, Kloppel G, Altmannsberger M. 1986. Differential diagnosis of gastrointestinal carcinomas by using monoclonal antibodies specific for individual keratin polypeptides. Lab Invest 55:497–504.

Padiath QS, Saigoh K, Schiffmann R, Asahara H, Yamada T, Koeppen A, Hogan K, Ptacek LJ, Fu YH. 2007. Corrigendum: lamin B1 duplications cause autosomal dominant leukodystrophy. Nat Genet 39:276.

Porter RM, Lane EB. 2003. Phenotypes, genotypes and their contribution to understanding keratin function. Trends Genet 19:278–285.

Pruitt KD, Tatusova T, Maglott DR. 2007. NCBI reference sequences (RefSeq): a curated non-redundant sequence database of genomes, transcripts and proteins. Nucleic Acids Res 35(Database issue): D61–D65.

Ramachandran RD, Perumalsamy V, Hejtmancik JF. 2007. Autosomal recessive juvenile onset cataract associated with mutation in BFSP1. Hum Genet 121:475–482.

Ramaekers FC, Vroom TM, Moesker O, Kant A, Scholte G, Vooijs GP. 1985. The use of antibodies to intermediate filament proteins in the differential diagnosis of lymphoma versus metastatic carcinoma. Histochem J 17:57–70.

Rogers MA, Winter H, Langbein L, Bleiler R, Schweizer J. 2004. The human type I keratin gene family: characterization of new hair follicle specific members and evaluation of the chromosome 17q212 gene domain. Differentiation 72:527–540.

Rogers MA, Edler L, Winter H, Langbein L, Beckmann I, Schweizer J. 2005. Characterization of new members of the human type II keratin

gene family and a general evaluation of the keratin gene domain on chromosome 12q1313. J Invest Dermatol 124:536–544.

Rothnagel JA, Traupe H, Wojcik S, Huber M, Hohl D, Pittelkow MR, Saeki H, Ishibashi Y, Roop DR. 1994. Mutations in the rod domain of keratin 2e in patients with ichthyosis bullosa of Siemens. Nat Genet 7:485–490.

Schweizer J, Bowden PE, Coulombe PA, Langbein L, Lane EB, Magin TM, Maltais L, Omary MB, Parry DA, Rogers MA, Wright MW. 2006. New consensus nomenclature for mammalian keratins. J Cell Biol 174:169–174.

Shackleton S, Lloyd DJ, Jackson SN, Evans R, Niermeijer MF, Singh BM, Schmidt H, Brabant G, Kumar S, Durrington PN, Gregory S, O'Rahilly S, Trembath RC. 2000. LMNA, encoding lamin A/C, is mutated in partial lipodystrophy. Nat Genet 24:153–156.

Speckman RA, Garg A, Du F, Bennett L, Veile R, Arioglu E, Taylor SI, Lovett M, Bowcock AM. 2000. Mutational and haplotype analyses of families with familial partial lipodystrophy (Dunnigan variety) reveal recurrent missense mutations in the globular C-terminal domain of lamin A/C. Am J Hum Genet 66:1192–1198.

Steinert PM, Parry DA. 1985. Intermediate filaments: conformity and diversity of expression and structure. Annu Rev Cell Biol 1:41–65.

Strelkov SV, Herrmann H, Geisler N, Wedig T, Zimbelmann R, Aebi U, Burkhard P. 2002. Conserved segments 1A and 2B of the intermediate filament dimer: their atomic structures and role in filament assembly. EMBO J 21:1255–1266.

Thompson JD, Higgins DG, Gibson TJ. 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res 22:4673–4680.

Vigouroux C, Magré J, Vantyghem MC, Bourut C, Lascols O, Shackleton S, Lloyd DJ, Guerci B, Padova G, Valensi P, Grimaldi A, Piquemal R, Touraine P, Trembath RC, Capeau J. 2000. Lamin A/C gene: sex-determined expression of mutations in Dunnigan-type familial partial lipodystrophy and absence of coding mutations in congenital and acquired generalized lipoatrophy. Diabetes 49: 1958–1962.

Wu KC, Bryan JT, Morasso MI, Jang SI, Lee JH, Yang JM, Marekov LN, Parry DA, Steinert PM. 2000. Coiled-coil trigger motifs in the 1B and 2B rod domain segments are required for the stability of keratin intermediate filaments. Mol Biol Cell 11:3539–3558.